Federated stochastic control of numerous heterogeneous energy storage systems

Emmanuel Gobet, Maxime Grangereau

Federated stochastic control of numerous heterogeneous energy storage systems *

Emmanuel GOBET † and Maxime GRANGEREAU ‡

Abstract

We propose a stochastic control problem to control cooperatively Thermostatically Controlled Loads (TCLs) to promote power balance in electricity networks. We develop a method to solve this stochastic control problem with a decentralized architecture, in order to respect privacy of individual users and to reduce both the telecommunications and the computational burden compared to the setting of an omniscient central planner. This paradigm is called federated learning in the machine learning community, see [YFY20], therefore we refer to this problem as a *federated stochastic control problem*. The optimality conditions are expressed in the form of a high-dimensional Forward-Backward Stochastic Differential Equation (FBSDE), which is decomposed into smaller FBSDEs modeling the optimal behaviors of the aggregate population of TCLs of individual agents. In particular, we show that these FBSDEs fully characterize the Nash equilibrium of a stochastic Stackelberg differential game. In this game, a coordinator (the leader) aims at controlling the aggregate behavior of the population, by sending appropriate signals, and agents (the followers) respond to this signal by optimizing their storage system locally. A mean-field-type approximation is proposed to circumvent telecommunication constraints and privacy issues. Convergence results and error bounds are obtained for this approximation depending on the size of the population of TCLs. A numerical illustration is provided to show the interest of the control scheme and to exhibit the convergence of the approximation. An implementation which answers practical industrial challenges to deploy such a scheme is presented and discussed.

1 Introduction

Context. To meet the goal of low carbon footprint for mitigating the climate change, the energy sector is seeking solutions for better energy-efficiency. Among them, the use of renewable energy (like solar or wind power) is appealing but on the other hand, the intermittency of their production raises challenges to satisfy the power balance between production and consumption. In this work, we focus on demand-side flexibilities, more specifically, Thermostatically Control Loads (TCLs) like for instance, fridges, air conditioners, hot water tanks, swimming pool heaters... These devices aim at maintaining a set-point temperature, but the realized temperature X has an inertia and tolerates a range of admissible values, which gives flexibility in controlling the appliances [BM16]. Leveraging this flexibility to provide services to the grid has an enormous potential [Mat+12; Cam+18].

Statement of the problem and objectives. In this paper, we consider the problem of optimally controlling a large population of N TCLs owned by individual consumers (also called agents), in a stochastic environment modeled by a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. From the application perspective, it corresponds to a centralized control architecture, where an omniscient planner solves a high-dimensional control problem in order to both minimize operational costs and promote energy balance.
The model will incorporate a common weather noise for all agents. This weather noise models the exogenous

*This work has benefited from several supports: Siebel Energy Institute (Calls for Proposals #2, 2016), ANR project CAESARS (ANR-15-CE05-0024), Association Nationale de la Recherche Technique (ANRT) and Electricité De France (EDF), *Finance for Energy Market* (FiME) Lab (Institut Europlace de Finance), Chair "Stress Test, RISK Management and Financial Steering" of the Foundation Ecole Polytechnique.

†Email: emmanuel.gobet@polytechnique.edu. Centre de Mathématiques Appliquées (CMAP), CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, Route de Saclay, 91128 Palaiseau Cedex, France. Corresponding author.

‡Email: maxime.grangereau@edf.fr. Electricité de France, Department OSIRIS, 7 Boulevard Gaspard Monge, 91120 Palaiseau cedex, France.

conditions that impact both the solar production through the irradiance [Bad+18] and the household consumption through non-constant lighting, heating, cooling, see [PMV02]. Conditionally to the common weather noise, the net consumption (household consumption without flexible appliance minus solar production) of different agents will be assumed independent. This allows us to account for spatial correlations of meteorological conditions [ADS99; ZDK16] between different agent locations, for instance.
To allow for general and realistic situations, the agents will differ w.r.t. their flexibilities and their consumption/production (difference of size and habits of the households, of renewable energy equipment, of appliances, etc). Additionally, we will assume that the agents are split in M classes in which the agents' flexibilities share the same physical characteristics (similar appliance), see Section 2 for the precise modeling. The net consumption of the i-th agent in the k-th class is denoted by $(\mathtt{P}^{\mathtt{load},(k,i)})_{k\in[M],i\in[N_k]}$ using the notation $[n] := \{1,\dots,n\}$ for any integer $n \geq 1$. The power consumed by its flexible appliance is denoted by $u^{(k,i,N)}$; hence, its total consumption is $\mathtt{P}^{\mathtt{load},(k,i)} + u^{(k,i,N)}$. Given a average (per agent) power $\mathtt{P}^{\mathtt{prod}}$ available on the public grid, the average power imbalance is $\frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}(u^{(l,j,N)} + \mathtt{P}^{\mathtt{load},(l,j)}) - \mathtt{P}^{\mathtt{prod}}$. Although not mathematically essential, $\mathtt{P}^{\mathtt{prod}}$ is assumed deterministic for simplicity, which makes sense since it comes from power demand forecast made by the planner (and usually supplied by conventional generation units). In addition, all the above quantities depend on time.
The optimization criterion will consist in minimizing the average power imbalance over a finite interval $[0,T]$ ($T > 0$ fixed), while maintaining each flexible power $u^{(k,i,N)}$ around a nominal value $u^{\mathtt{ref},(k,i)}$ and the temperature $X^{(k,i,N)}$ around a set-point temperature $x^{\mathtt{ref},(k,i)}$. All in all, the controls are $(u^{(k,i,N)})_{k\in[M],i\in[N_k]}$, adapted to the ambient filtration $\mathbb{F}$, and the cost functional takes the form

$$\mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k}\left\{\int_0^T\left(\frac{\mu_t^{(k)}}{2}\left(u_t^{(k,i,N)} - u_t^{\mathtt{ref},(k,i)}\right)^2 + \frac{\nu_t^{(k)}}{2}\left(X_t^{(k,i,N)} - x_t^{\mathtt{ref},(k,i)}\right)^2\right)\mathrm{d}t + \frac{\rho^{(k)}}{2}\left(X_T^{(k,i,N)} - x_T^{\mathtt{f},(k,i)}\right)^2\right\}\right]$$
$$+\,\mathbb{E}\left[\int_0^T \mathcal{L}_t\left(\frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}(u_t^{(l,j,N)} + \mathtt{P}_t^{\mathtt{load},(l,j)}) - \mathtt{P}_t^{\mathtt{prod}}\right)\mathrm{d}t\right]. \tag{1.1}$$

The fact that the parameters $\mu^{(k)}$, $\nu^{(k)}$, $\rho^{(k)}$ may depend on the class of device allows to model heterogeneity among the devices. For instance, a fridge may not have the same temperature dead-band tolerance as a heat pump, which justifies to consider different values of the parameter $\nu^{(k)}$ for these two classes of devices. Actually, we are looking for a solution method where agents can keep their individual data private. This privacy preservation is nowadays a major concern in grid management, see [AA19] for a recent overview and references therein. This is quite topical in AI systems, see [BG19]. We are also interested in deriving a practical implementation of the control where minimal communication between actors is required. All these concerns of data privacy, managing heterogeneous agents, low communication exchanges are refereed to as federated learning in the machine learning community [YFY20]. In the stochastic control community, it is seemingly new and we believe it will be increasingly important from the mathematical modelling point of view in the next years.

Methodology and main contributions. Mathematically speaking, our problem fits the setting of stochastic control problem in high dimension (the number N of agents). The dynamics for each state variable $X^{(k,i,N)}$ (modeling the temperature inertia) is linear w.r.t. the control (appliance power): for each agent, it typically writes under the form[1] [2]

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \underbrace{u_t/C}_{\text{appliance power consumption}} - \underbrace{(X_t - X_t^{out})/RC}_{\text{thermal losses}} + \underbrace{\Lambda_t}_{\text{exogenous perturbations}}. \tag{1.2}$$

The filtration can be quite general and allows, for instance, for jump processes in the net consumption. The stochastic (Pontryagin) maximum principle enables us to characterize the optimal controls as solution to a coupled system of Forward-Backward Stochastic Differential Equations (FBSDE), see Theorem 2.1. However, the FBSDE is a high-dimensional coupled equation, and hence curse of dimensionality occurs.

[1] the coefficient C is the calorific capacity of the system C, R is the thermal resistance of the system, X^{out} is the temperature of the environment, and Λ models random perturbations like opening the fridge, using the hot water tank for a shower etc

[2] the affine-linear dynamic is in agreement with the "leaky battery model" presented in [Hao+14; TTS16] and with the first-order dynamical model used to model the temperature evolution of a TCL in [DP+19].

To overcome this, we design a decoupled system, with a so-called coordination problem and individual problems associated to each agent. The coordination problem is a FBSDE which can be interpreted as the optimality conditions of a (convex) optimal control problem that a coordinator has to solve to compute a coordination signal. Each individual problem is an FBSDE which can be interpreted as the optimality conditions of the control problem of a selfish agent controlling (with locally available information) its individual storage system to minimize operational costs, while responding to the coordination signal. In other words, we show that the optimal solution of the control problem of the central planner corresponds to the (unique) Nash equilibrium of a stochastic Stackelberg differential game, which allows for possible decentralized control schemes. This is in some way the inverse perspective of potential stochastic differential games [FMHL19], in which one seeks a stochastic control problem which optimality conditions coincide with the Nash system of a given stochastic differential game. In order to avoid the need for real-time communication from agents to the coordinator, we design a mean-field-type approximation of the decoupled system: the approximation of the coordination problem mainly depends on the population statistics and not anymore on the data of agents. Under the assumption that the size of each class of agents is large enough and under the assumption that agents are conditionally independent in some sense, we prove that this new decoupled system yields a control which preserves privacy and converges to the omniscient control in the limit of infinite population (see Theorem 4.12). Error bounds on performance loss are also derived. Besides, there again, we can interpret the approximations of the coordination and individual problems as the system of FBSDE characterizing the (unique) equilibrium of a stochastic Stackelberg differential game with a leader, the coordinator, and many (non-symmetric) followers, the agents. To get the convergence results and error bounds, we leverage the conditional Law of Large Numbers, stability results for FBSDEs and probabilistic properties related to immersion of filtration (to deal with the common noise). The new control boils down to solving a system of $M + 1 \ll N$ weakly coupled FBSDEs, and thus one suffers much less from the curse of dimensionality than what may be expected. We illustrate these results numerically on an example involving a large population of two types of devices (water heaters and heat pumps) with realistic characteristics. In this work, we also discuss a practical online and decentralized implementation of the privacy-preserving control, see Algorithm 1. It allows real-time computations of a coordination signal by a coordinator. Broadcasting this signal allows each agent to compute its optimal response in real-time. In particular, the coordination signal sent at each time t to all agents is a function of time measurable with respect to the information available to the coordinator at time t. Besides, the problems solved online by the agents are easy to solve and only require information available locally. This allows to preserve the privacy of individual consumers and maintain quality of service if communication loss occurs.

Literature background. Stochastic control of large population of micro-grids has recently drawn significant interest, but standard methods like Stochastic Dynamic Programming, Stochastic Dual Dynamic Programming (in a discrete-time setting) suffer rapidly from the curse of dimensionality, when considering more than a few tens of micro-grids, even when considering spatial decomposition techniques [Car+19; Car+20]. To tackle this issue, mean-field approximations are particularly promising, as they become more accurate when the number of agents grows. There is a recent and abundant literature about stochastic control with large population, commonly known as mean-field games (MFG)/McKean-Vlasov (MKV) stochastic control problems, see [CDL13; CD18; BFY+13] among recent contributions. Mean-Field Games models for control of large populations of micro-grids without common noise have been proposed in [DP+19; KM13; KM16]. In [DP+19], self-interested consumers allocate their flexible consumption and choose a level of participation to electricity reserves mechanisms according to price signals derived from a Unit-Commitment problem solved by a coordinator. In [KM13], water heaters are controlled so that their average profile tracks a specific profile sent by a coordinator. This model is enriched in [KM16] to consider Markovian jumps dynamics for individual water heaters and non-uniformity of the temperature within water tanks.

In the literature of MFG and MKV control, agents are usually assumed to be symmetric. Let us mention however [HMC+06; KM13; KM16; ATM20], [BFY+13, Chapter 8, pp. 67-72] which consider an heterogeneous population by introducing user classes in the setting of Mean-Field Games.

Our model is defined with a common noise, which seemingly connects our contribution to the recent developments of the theory of MFG/MKV with a common noise. In [ATM20], similarly as in our work, a setting with common and individual noises is considered, heterogeneity among agents is introduced, and the structure of equations obtained is similar. However, this work considers applications related to price-arbitrage and peak-shaving, and directly studies the Mean-Field approximation. By contrast, our work focuses on tracking power imbalance from an application point of view, and we are also interested in approximation and convergence results. The model of [ATM20] is extended to the case with jumps in [MMS19]. These works mostly consider the case of an infinite population, whereas we consider a finite number of agents.

Note that Mean-Field Games assume a competitive setting and seek for Nash equilibria, which may be far from optimal from a collective point of view. This performance loss can easily be assessed in Linear-Quadratic frameworks, as both MKV control and MFG admit explicit feedback formulas for the control, which allows to compute the Price of Anarchy. In this paper, by contrast to the MFG frameworks for the control of micro-grids [DP+19; KM13; KM16; ATM20; MMS19], we assume a cooperative setting, so that our problem is closely related to the field of MKV stochastic control problems. We show that the optimal solution of our control problem is also the (unique) Nash equilibrium of a Stochastic Stackelberg Differential Game with a leader (the coordinator) and many heterogeneous followers (the agents), whose decisions are impacted by the decisions of the leader. Hence, we follow the inverse perspective of potential stochastic differential games [FMHL19], in which one seeks a stochastic control problem whose optimality conditions coincide with the Nash system of a given stochastic game. See [ŞC14] or [CD18, section 7.1, pp. 541-610, Volume II] for an introduction to Mean-Field Games with major and minor players.

We also mention other works presenting cooperative control architectures of TCLs, without a priori optimality guarantees. [Hao+13; Hao+14] propose a control architecture based on priority queue to decide which device to control. A control architecture based on PDE models is proposed in [TTS15] to track a reference profile where each device builds an ensemble model for the whole population and uses this model to compute an appropriate random switching rate. Another control architecture is proposed in [Tro+16] which ensures that the aggregate consumption of a large population of TCLs depends linearly on the frequency of the network (which is a good indicator for power balance) and its rate of change. Our approach differs from these works since we use tools coming from stochastic optimal control theory, which provides a priori optimality guarantee (up to model errors).

Organization of the paper. The probabilistic model as well as first-order necessary and sufficient optimality conditions (Theorem 2.1) are given in Section 2. A decomposition method for the optimality system is obtained in Section 3, with definitions of the coordination problem (Proposition 3.1) and individual problems (Proposition 3.3). Then approximations of these problems are given in Section 4. In particular, the solution of the approximation of the coordination problem is shown to be progressively measurable with respect to the common noise filtration, see Theorem 4.5, which is a desirable property in a decentralized control scheme. Indeed it allows a third party (called coordinator) to solve the approximate coordination problem without having to observe the aggregated individual parameters, circumventing privacy and telecommunication issues. Error bounds between the privacy-preserving control and the omniscient control are then presented, see Theorem 4.12. Section 5 collects a few numerical illustrations in the case of agents equipped with heat pumps or water heaters. Practical interest of the approach is demonstrated, as well as the convergence of the mean-field approximation in the limit of large populations. Then, a decentralized online implementation with minimal information sharing of the approximate solution method for the control problem is presented in Section 6. Some of the proofs are postponed to Section 7.

Most commonly used notations. We list the most common notations used in this paper.

$\triangleright$ *Numbers, vectors, matrices.* $\mathbb{R}$, $\mathbb{N}$, $\mathbb{N}^*$ denote respectively the set of real numbers, integers and positive integers. The notation $|x|$ stands for the Euclidean norm of a vector x. For $k \in \mathbb{N}$, the notation $[k]$ stands for the integer set $\{1, ..., k\}$.

$\triangleright$ *Function derivatives.* For a smooth function $g : \mathbb{R}^p \mapsto \mathbb{R}$, g'_x represents the partial derivative of g with respect to x. However, the notation x_t refers to the value of a process x at time t (and not to the partial derivative of x with respect to t).

$\triangleright$ *Probability.* The randomness on the interval $[0, T]$ is modeled on a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, with a right-continuous filtration $\mathbb{F} := \{\mathcal{F}_t\}_{0\leq t\leq T}$ augmented with the $\mathbb{P}$-null sets. We consider another filtration $\mathbb{G} \subset \mathbb{F}$ (meaning $\mathcal{G}_t \subset \mathcal{F}_t$ for all t in $[0, T]$), assumed immersed in $\mathbb{F}$ (see [CD18, Definition 1.2, p.5, Volume II]: all $\mathbb{G}$ square integrable martingale are $\mathbb{F}$-martingales): this filtration will model the structure of information for the common weather noise. The immersion property implies the independence of $\mathcal{G}_T$ and $\mathcal{F}_t$ conditionally on $\mathcal{G}_t$, for any $t \in [0, T]$.

This assumption is equivalent to the fact that for any $t \in [0,T]$, for any random variable $X \in \mathbb{L}^1(\mathcal{F}_t), \mathbb{E}[X|\mathcal{G}_T] = \mathbb{E}[X|\mathcal{G}_t]$ (see [CD18, Proposition 1.3, p. 6, Volume II]).

The set of square integrable variables is denoted by $\mathbb{L}^2$. The notation $\mathbb{L}^2_T$ stands for the set of $\mathcal{F}_T$-measurable square integrable variables.

$\triangleright$ *Stochastic processes.* For a vector/matrix-valued random variable V, its conditional expectation with respect to the sigma-field $\mathcal{F}_t$ is denoted by $\mathbb{E}_t[Z] = \mathbb{E}[Z|\mathcal{F}_t]$.

All the martingales are considered with their càdlàg modifications.

The space $\mathcal{S}$ (resp. $\mathcal{H}$) stands for the $\mathbb{F}$-adapted càdlàg (resp. $\mathbb{F}$-progressively measurable) processes $(\Psi_t : t \in [0,T])$ valued in an Euclidean space $\mathcal{E}$ such that $\sqrt{\mathbb{E}\left[\sup_{t\in[0,T]} |\Psi_t|^2\right]} =: \|\Psi\|_{\mathcal{S}}$ (resp. $\sqrt{\mathbb{E}\left[\int_0^T |\Psi_t|^2 \mathrm{d}t\right]} =: \|\Psi\|_{\mathcal{H}}$). Since the space $\mathcal{E}$ will be clear from the context (typically $\mathbb{R}$, $\mathbb{R}^M$ or $\mathbb{R}^N$), we will skip the reference to it in the notation. The space $\mathcal{H}_{\mathbb{G}}$ (resp. $\mathcal{S}_{\mathbb{G}}$) is the subspace of processes in $\mathcal{H}$ (resp. $\mathcal{S}$) which are $\mathbb{G}$-progressively measurable.

2 Model, assumptions and first properties

2.1 Assumptions

We follow the model presented in introduction with $N = \sum_{k=1}^M N_k$ agents split into M classes of N_k agents each (see Fig. 1a[3]). The state variable (temperature) for the i-th agent of the k-th class is $X^{(k,i,N)}$ and satisfies to the dynamics

$$X_t^{(k,i,N)} = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s^{(k,i,N)} + \beta_s^{(k)} X_s^{(k,i,N)} + \gamma_s^{(k,i)}\right) \mathrm{d}s, \tag{2.1}$$

where $u^{(k,i,N)}$ is the control for the flexible appliance of agent (k,i). The above dynamics is consistent with the example in (1.2), in particular the coefficients $\alpha^{(k)}$ and $\beta^{(k)}$ are the same within the class k (similar device). This allows to incorporate heterogeneity for the devices, which do not have the same performances nor thermal behaviors. On the mathematical side, we assume from now on that

(H-X) $\alpha^{(k)}, \beta^{(k)}$ are measurable deterministic functions, uniformly bounded on $[0,T]$. Each process $\gamma^{(k,i)}$ is in $\mathcal{H}$. Each $x_0^{(k,i)}$ is deterministic.

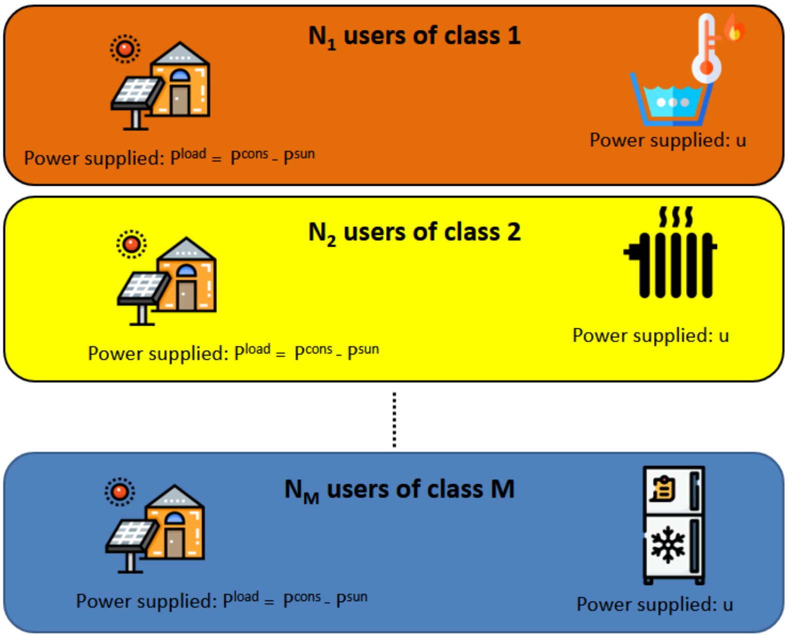
Given the control $u^{(N)} := (u^{(k,i,N)})_{k,i} \in \mathcal{H}$, the functional to minimize is

$$\begin{aligned}\mathcal{J}(u) =& \mathbb{E}\left[\frac{1}{N}\sum_{k=1}^M \sum_{i=1}^{N_k} \left\{\int_0^T \left(\frac{\mu_t^{(k)}}{2}\left(u_t^{(k,i,N)} - u_t^{\mathtt{ref},(k,i)}\right)^2 + \frac{\nu_t^{(k)}}{2}\left(X_t^{(k,i,N)} - x_t^{\mathtt{ref},(k,i)}\right)^2\right)\mathrm{d}t + \frac{\rho^{(k)}}{2}\left(X_T^{(k,i,N)} - x_T^{\mathtt{f},(k,i)}\right)^2\right\}\right] \\ &+ \mathbb{E}\left[\int_0^T \mathcal{L}_t\left(\frac{1}{N}\sum_{l=1}^M \sum_{j=1}^{N_l} (u_t^{(l,j,N)} + \mathrm{P}_t^{\mathtt{load},(l,j)}) - \mathrm{P}_t^{\mathtt{prod}}\right)\mathrm{d}t\right],\end{aligned} \tag{2.2}$$

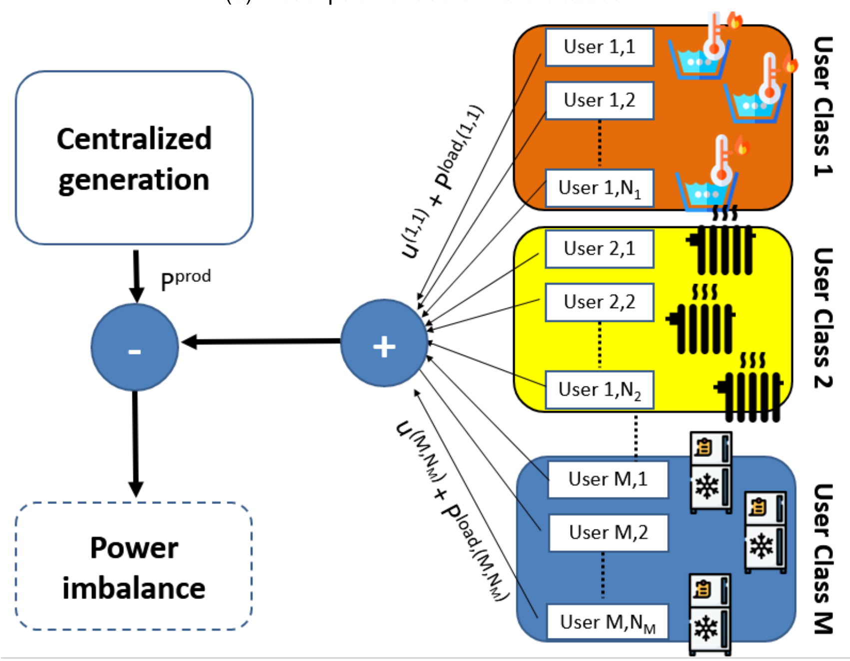
corresponding to an omniscient planner aiming to control TCLs to track the power imbalance signal, represented in Figure 1b, while keeping each individual flexibility around a possibly stochastic nominal state (associated to $u^{\mathtt{ref},(k,i)}$, $x^{\mathtt{ref},(k,i)}$, $x_T^{\mathtt{f},(k,i)}$).

(H-J) $\mathrm{P}^{\mathtt{prod}}$ is a measurable deterministic function, square integrable on $[0,T]$.
All the processes $u^{\mathtt{ref},(k,i)}, x^{\mathtt{ref},(k,i)}, \mathrm{P}^{\mathtt{load},(k,i)}$ are in $\mathcal{H}$ and $x_T^{\mathtt{f},(k,i)}$ is in $\mathbb{L}^2_T$.
The coefficients $\mu^{(k)}, \nu^{(k)}$ are deterministic measurable functions, the $\rho^{(k)}$ are deterministic. They are all bounded. In addition, for some $\varepsilon > 0$, we have $\mu_t^{(k)} \geq \varepsilon$ and $\nu_t^{(k)}$ for any t and k. The function $\mathcal{L} : (t,x) \in [0,T] \times \mathbb{R} \mapsto \mathbb{R}$ is deterministic measurable. We assume that for any $t \in [0,T]$, $x \mapsto \mathcal{L}_t(x) = \mathcal{L}(t,x)$ is convex, twice continuously differentiable, with uniformly bounded second order derivative.

[3] Icons made by Freepik and Smashicons from www.flaticon.com



(a) Description of users in the classes



(b) Power balance

Figure 1: Heterogeneity of agents and power imbalance

2.2 Differentiability, convexity, characterization of optimality

We start with a somehow standard result. We show that, under our assumptions, $\mathcal{J}$ is strongly convex and admits a unique minimizer which can be obtained by solving a Forward-Backward Stochastic Differential Equation, obtained using the Stochastic Pontryagin Principle. The proof is postponed to Subsection 7.1.

Theorem 2.1. *The function* $\mathcal{J} : \mathcal{H} \mapsto \mathbb{R}$ *is strongly convex. It admits a unique minimizer denoted* $u^{(N)} :=$

$(u^{(k,i,N)})_{k\in[M],i\in[N_k]} \in \mathcal{H}$. *Define* $X^{(N)} = (X^{(k,i,N)})_{k\in[M],i\in[N_k]} \in \mathcal{S}$ *by* (2.1) *and* $Y^{(N)} = (Y^{(k,i,N)})_{k\in[M],i\in[N_k]} \in \mathcal{S}$ *by:*

$$Y_t^{(k,i,N)} = \mathbb{E}_t\left[\rho^{(k)}(X_T^{(k,i,N)} - x_T^{\mathtt{f},(k,i)}) + \int_t^T \left(\beta_s^{(k)} Y_s^{(k,i,N)} + \nu_s^{(k)}(X_s^{(k,i,N)} - x_s^{\mathtt{ref},(k,i)})\right) \mathrm{d}s\right]. \tag{2.3}$$

Then $(u^{(k,i,N)}, X^{(k,i,N)}, Y^{(k,i,N)})_{k\in[M],i\in[N_k]} \in \mathcal{H}\times\mathcal{S}\times\mathcal{S}$ *is the unique solution in* $\mathcal{H}\times\mathcal{S}\times\mathcal{S}$ *of the coupled FBSDE with unknowns* $(u^{(k,i)}, X^{(k,i)}, Y^{(k,i)})_{k\in[M],i\in[N_k]} \in \mathcal{H}\times\mathcal{S}\times\mathcal{S}$*:*

$$\begin{aligned}
&\forall t \in [0,T], \forall k \in [M], \forall i \in [N_k],\\
&\begin{cases}
X_t^{(k,i)} = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s^{(k,i)} + \beta_s^{(k)} X_s^{(k,i)} + \gamma_s^{(k,i)}\right) \mathrm{d}s,\\
Y_t^{(k,i)} = \mathbb{E}_t\left[\rho^{(k)}\left(X_T^{(k,i)} - x_T^{\mathtt{f},(k,i)}\right) + \int_t^T \left(\beta_s^{(k)} Y_s^{(k,i)} + \nu_s^{(k)}\left(X_s^{(k,i)} - x_s^{\mathtt{ref},(k,i)}\right)\right) \mathrm{d}s\right],\\
\mu_t^{(k)}\left(u_t^{(k,i)} - u_t^{\mathtt{ref},(k,i)}\right) + \mathcal{L}_x'\left(t, \frac{1}{N}\sum_{l=1}^M \sum_{j=1}^{N_l} (u_t^{(l,j)} + \mathrm{P}_t^{\mathtt{load},(l,j)}) - \mathrm{P}_t^{\mathtt{prod}}\right) + \alpha_t^{(k)} Y_t^{(k,i)} = 0.
\end{cases}
\end{aligned} \tag{2.4}$$

Our assumptions guarantee that the optimal control problem

$$\begin{aligned}
\min_{u\in\mathcal{H}} \quad & \mathcal{J}(u)\\
s.t. \quad & (2.1)
\end{aligned}$$

has a unique solution $u^{(N)}$, which can be equivalently computed by solving (2.4). This system is a high-dimensional coupled FBSDE, the dimension being the number N of agents. Hence, it suffers from the curse of dimensionality, as the number of agents may be very large. To tackle this issue, we propose a decomposition method of the FBSDE (2.4) in the next section.

3 Decomposition of the problem and equivalent representation as a stochastic differential game

Notations

- *Empirical and statistical means:* We introduce the notation $\bar{\mathrm{P}}^{\mathtt{load},(N)} := \frac{1}{N}\sum_{l=1}^M \sum_{j=1}^{N_l} \mathrm{P}^{\mathtt{load},(l,j)}$ for the empirical mean process of net consumption of agents.
- *Empirical means over a class:* We introduce the following notations for the empirical means: $\bar{\gamma}^{(k,N)} := \frac{1}{N_k}\sum_{j=1}^{N_k} \gamma^{(k,j)}$, $\bar{u}^{\mathtt{ref},(k,N)} := \frac{1}{N_k}\sum_{j=1}^{N_k} u^{\mathtt{ref},(k,j)}$, $\bar{x}^{\mathtt{ref},(k,N)} := \frac{1}{N_k}\sum_{j=1}^{N_k} x^{\mathtt{ref},(k,j)}$, $\bar{x}_T^{\mathtt{f},(k,N)} := \frac{1}{N_k}\sum_{j=1}^{N_k} x_T^{\mathtt{f},(k,j)}$.

In this section, we show that the control problem is equivalent to two types of control problems arising in a nested structure. We call the first problem the *coordination problem*, as it allows to compute a coordination signal. Once the coordination signal has been computed, the control problem can be decomposed into N sub-problems, the *individual problems*, each of them associated to an individual consumer/agent. The parameters of the individual problem of each agent only involve the individual data of the corresponding agent (consumption and preferences), the shared information $\mathbb{G}$ and the coordination signal. We also show that the coordination problem (resp. each individual problem) can be interpreted as the optimality conditions of a control problem of the coordinator (resp. of each agent). This shows that the optimal solution of the control problem corresponds to the (unique) Nash equilibrium of a stochastic differential game, allowing for a decentralized implementation.

3.1 The coordination problem

Proposition 3.1. *Consider the empirical mean processes for all* $k \in [M]$*:*

$$(\bar{u}^{(k,N)}, \bar{X}^{(k,N)}, \bar{Y}^{(k,N)}) := \left(\frac{1}{N_k}\sum_{j=1}^{N_k} u^{(k,j,N)}, \frac{1}{N_k}\sum_{j=1}^{N_k} X^{(k,j,N)}, \frac{1}{N_k}\sum_{j=1}^{N_k} Y^{(k,j,N)}\right) \in \mathcal{H}\times\mathcal{S}\times\mathcal{S}.$$

The empirical mean process $(\bar{u}^{(k,N)}, \bar{X}^{(k,N)}, \bar{Y}^{(k,N)})_{k\in[M]}$ is the unique solution of the following FBSDE with unknowns $(u^{(k)}, X^{(k)}, Y^{(k)})_{k\in[M]} \in \mathcal{H}\times\mathcal{S}\times\mathcal{S}$, which we call the coordination problem:

$$\forall t \in [0,T], \forall k \in [M],$$
$$\begin{cases} X_t^{(k)} = \frac{1}{N_k}\sum_{j=1}^{N_k} x_0^{(k,j)} + \int_0^t \left(\alpha_s^{(k)} u_s^{(k)} + \beta_s^{(k)} X_s^{(k)} + \bar{\gamma}_s^{(k,N)}\right) \mathrm{d}s, \\ Y_t^{(k)} = \mathbb{E}_t\left[\rho^{(k)}\left(X_T^{(k)} - \bar{x}_T^{\mathtt{f},(k,N)}\right) + \int_t^T \left(\beta_s^{(k)} Y_s^{(k)} + \nu_s^{(k)}\left(X_s^{(k)} - \bar{x}_s^{\mathtt{ref},(k,N)}\right)\right) \mathrm{d}s\right], \\ \mu_t^{(k)}\left(u_t^{(k)} - \bar{u}_t^{\mathtt{ref},(k,N)}\right) + \mathcal{L}_x'\left(t, \sum_{l=1}^M \pi^{(l)} u_t^{(l)} + \bar{\mathrm{P}}_t^{\mathtt{load},(N)} - \mathrm{P}_t^{\mathtt{prod}}\right) + \alpha_t^{(k)} Y_t^{(k)} = 0. \end{cases} \tag{3.1}$$

The proof of the above Proposition is postponed to Subsection 7.2. This proof shows that the FBSDE (3.1) is the optimality system of the stochastic control problem (7.3), which can be interpreted as the control problem of a coordinator aiming at controlling the aggregate behaviors of the agents within different classes.

Definition 3.2. *Let $(\bar{u}^{(k,N)}, \bar{X}^{(k,N)}, \bar{Y}^{(k,N)})_{k\in[M]} \in \mathcal{H}\times\mathcal{S}\times\mathcal{S}$ be the unique solution of the coordination problem* (3.1). *We define the coordination signal $\bar{v}^{(N)} \in \mathcal{H}$ by:*

$$\forall t \in [0,T], \quad \bar{v}_t^{(N)} := \mathcal{L}_x'\left(t, \sum_{l=1}^M \pi^{(l)} \bar{u}_t^{(l,N)} + \bar{\mathrm{P}}_t^{\mathtt{load},(N)} - \mathrm{P}_t^{\mathtt{prod}}\right). \tag{3.2}$$

3.2 The individual problems

Proposition 3.3. *Let $\bar{v}^{(N)} \in \mathcal{H}$ be the coordination signal defined in* (3.2). *For any $k \in [M], i \in [N_k]$, consider the FBSDE with unknown $(u, X, Y) \in \mathcal{H}\times\mathcal{S}\times\mathcal{S}$, called individual problem of agent i of class k:*

$$\forall t \in [0,T],$$
$$\begin{cases} X_t = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s + \beta_s^{(k)} X_s + \gamma_s^{(k,i)}\right) \mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\rho^{(k)}\left(X_T - x_T^{\mathtt{f},(k,i)}\right) + \int_t^T \left(\beta_s^{(k)} Y_s + \nu_s^{(k)}\left(X_s - x_s^{\mathtt{ref},(k,i)}\right)\right) \mathrm{d}s\right], \\ \mu_t^{(k)}\left(u_t - u_t^{\mathtt{ref},(k,i)}\right) + \bar{v}_t^{(N)} + \alpha_t^{(k)} Y_t = 0. \end{cases} \tag{3.3}$$

Then (3.3) *has a unique solution $(u^{(k,i,N)}, X^{(k,i,N)}, Y^{(k,i,N)}) \in \mathcal{H}\times\mathcal{S}\times\mathcal{S}$. Besides $(u^{(k,i,N)}, X^{(k,i,N)}, Y^{(k,i,N)})_{k\in[M],i\in[N_k]}$ is the unique solution of the FBSDE* (2.4) *and $(u^{(k,i,N)})_{k\in[M],i\in[N_k]}$ is the unique solution of control problem $\min_{u^{(N)}\in\mathcal{H}} \mathcal{J}(u^{(N)})$.*

Proof. The FBSDE (3.3) fully characterizes the solutions of the following stochastic control problem:

$$\begin{aligned} \min_{u\in\mathcal{H}} \quad & \mathbb{E}\left[\int_0^T \left(\frac{\mu_t^{(k)}}{2}\left(u_t - u_t^{\mathtt{ref},(k,i)}\right)^2 + \frac{\nu_t^{(k)}}{2}\left(X_t - x_t^{\mathtt{ref},(k,i)}\right)^2 + \bar{v}_t^{(N)} u_t\right) \mathrm{d}t + \frac{\rho^{(k)}}{2}\left(X_T - x_T^{\mathtt{f},(k,i)}\right)^2\right], \\ s.t. \quad & X_t = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s + \beta_s^{(k)} X_s + \gamma_s^{(k,i)}\right) \mathrm{d}s. \end{aligned} \tag{3.4}$$

This control problem (3.4) has a unique solution, by similar arguments as in the proof of Proposition 3.1, hence so does the FBSDE (3.3). The fact that $(u^{(k,i,N)})_{k\in[M],i\in[N_k]}$ is the unique minimizer of $\mathcal{J}$ is a consequence of Proposition 3.1 and of the uniqueness of the solutions of (2.4), (3.1),(3.3). □

The stochastic control problem (3.4) can be interpreted as the control problem of agent i of class k interacting with an aggregator which sends him a coordination signal. It can be interpreted as the agent aiming at minimizing operational costs and a cost of contribution to global power imbalance, where the coordination signal $\bar{v}^{(N)}$ plays the role of a price signal cast by the coordinator. In particular, replacing $\bar{v}^{(N)}$ by its expression, the above Proposition shows that the optimal solution $(u^{(N)})$ of the control problem corresponds to the (unique) Nash equilibrium of a Stochastic differential game.

The following Proposition gives another way to compute the solution of the individual problems and allows to focus on the fluctuations of the controls and states of individual players around the empirical means over classes.

Proposition 3.4. *Let $(u^{(k,i,N)}, X^{(k,i,N)}, Y^{(k,i,N)}) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ be the unique solution of the individual problem of agent $i \in [N_k]$ of class $k \in [M]$. Let $(\bar{u}^{(k,N)}, \bar{X}^{(k,N)}, \bar{Y}^{(k,N)})_{k\in[M]} \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ be the unique solution of the coordination problem. Then $(u^{(k,i,N)} - \bar{u}^{(k,N)}, X^{(k,i,N)} - \bar{X}^{(k,N)}, Y^{(k,i,N)} - \bar{Y}^{(k,N)}) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ is the unique solution of the following FBSDE with unknowns $(\Delta u, \Delta X, \Delta Y) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$:*

$$\forall t \in [0,T],$$
$$\begin{cases} \Delta X_t = x_0^{(k,i)} - \bar{x}_0^{(k,N)} + \int_0^t \left(\alpha_s^{(k)} \Delta u_s + \beta_s^{(k)} \Delta X_s + \gamma_s^{(k,i)} - \bar{\gamma}_s^{(k,N)}\right) \mathrm{d}s, \\ \Delta Y_t = \mathbb{E}_t \left[\rho^{(k)} \left(\Delta X_T - x_T^{\mathtt{f},(k,i)} + \bar{x}_T^{\mathtt{f},(k,N)}\right) + \int_t^T \left(\beta_s^{(k)} \Delta Y_s + \nu_s^{(k)} \left(\Delta X_s - x_s^{\mathtt{ref},(k,i)} + \bar{x}_s^{\mathtt{ref},(k,N)}\right)\right) \mathrm{d}s\right], \\ \mu_t^{(k)} \left(\Delta u_t - u_t^{\mathtt{ref},(k,i)} + \bar{u}_t^{\mathtt{ref},(k,N)}\right) + \alpha_t^{(k)} \Delta Y_t = 0. \end{cases} \tag{3.5}$$

We have developed a decomposition method of the N-dimensional FBSDE (2.4), which shows that it is equivalent to solve one M-dimensional FBSDE, with $M << N$ typically, the coordination problem, and N one-dimensional FBSDE. However the parameters of the FBSDE depend on aggregate data of individual agents. In practical applications, the coordination problem shall be solved by a third party, called coordinator, which may not have access in real time to the aggregate data of individual agents (namely $\bar{\mathtt{P}}^{\mathtt{load},(k,N)}$, $\bar{\gamma}^{(k,N)}$, $\bar{u}^{\mathtt{ref},(k,N)}$, $\bar{x}^{\mathtt{ref},(k,N)}$, $\bar{x}_T^{\mathtt{f},(k,N)}$). The reason for that is two-fold: a privacy concern, as agents may not wish to share their private data (namely $\mathtt{P}^{\mathtt{load},(k,i)}$, $\gamma^{(k,i)}$, $u^{\mathtt{ref},(k,i)}$, $x^{\mathtt{ref},(k,i)}$, $x_T^{\mathtt{f},(k,i)}$), and a technical reason, as heavy telecommunication infrastructures would be required to allow individual agents to share their private data with the coordinator.

4 Approximate decentralized control architecture preserving privacy

In this section, we present an approximation of the coordination problem which parameters only depend on the statistical behaviors of the agents (namely the conditional means of their associated parameters), rather than the actual realization of their individual data. In particular, this is a desirable feature in decentralized implementations where the coordination problem is solved by a third party not observing in real-time the aggregate behaviors of agents, for privacy reasons or in order to reduce real-time telecommunication requirements. Solving the approximate coordination problem allows to compute an approximation of the coordination signal (3.2), which may be used to decouple individual problems. This is done at the cost of a small performance loss, which can be estimated and which vanishes asymptotically when the number of consumers/agents goes to infinity.

Notations

- *Empirical conditional means:* We introduce the notation $\bar{\mathtt{P}}_t^{\mathtt{load}} := \frac{1}{N} \sum_{l=1}^{M} \sum_{j=1}^{N_l} \mathbb{E}\left[\mathtt{P}_t^{\mathtt{load},(l,j)} | \mathcal{G}_t\right]$ for the conditional average of the empirical average consumption.
- *Empirical conditional means over a class:* We introduce the following notations for the conditional means of the empirical average over classes of individual parameters: for $t \in [0,T]$, $\bar{\gamma}_t^{(k)} := \frac{1}{N_k} \sum_{j=1}^{N_k} \mathbb{E}\left[\gamma_t^{(k,j)} | \mathcal{G}_t\right]$, $\bar{u}_t^{\mathtt{ref},(k)} := \frac{1}{N_k} \sum_{j=1}^{N_k} \mathbb{E}\left[u_t^{\mathtt{ref},(k,j)} | \mathcal{G}_t\right]$, $\bar{x}_t^{\mathtt{ref},(k)} := \frac{1}{N_k} \sum_{j=1}^{N_k} \mathbb{E}\left[x_t^{\mathtt{ref},(k,j)} | \mathcal{G}_t\right]$, $\bar{x}_T^{\mathtt{f},(k)} := \frac{1}{N_k} \sum_{j=1}^{N_k} \mathbb{E}\left[x_T^{\mathtt{f},(k,j)} | \mathcal{G}_T\right]$.

Lemma 4.1. *The processes $(\bar{\gamma}^{(k)}, \bar{u}^{\mathtt{ref},(k)}, \bar{x}^{\mathtt{ref},(k)}, \bar{x}_T^{\mathtt{f},(k)})_{k\in[M]}$, $\bar{\mathtt{P}}^{\mathtt{load}}$ are $\mathbb{G}$-measurable and can be assumed progressively measurable without loss of generality, and therefore in $\mathcal{H}_{\mathbb{G}}$.*

Proof. Let Z denote any of the processes $(\bar{\gamma}^{(k,N)}, \bar{u}^{\mathtt{ref},(k,N)}, \bar{x}^{\mathtt{ref},(k,N)}, \bar{x}_T^{\mathtt{f},(k,N)})_{k\in[M]}$, $\bar{\mathtt{P}}^{\mathtt{load},(N)}$. Then $Z \in \mathcal{H}$ is $\mathbb{F}$-progressively measurable and since $\mathbb{G}$ is immersed in $\mathbb{F}$:

$$\forall t \in [0,T], \quad \mathbb{E}\left[Z_t | \mathcal{G}_t\right] = \mathbb{E}\left[Z_t | \mathcal{G}_T\right].$$

As in [CD18, Volume II, pp. 265-266], one can redefine $t \mapsto \mathbb{E}\left[Z_t | \mathcal{G}_T\right]$ as a $\mathcal{B}([0,T]) \otimes \mathbb{G}$-measurable process, up to a $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}$-null set. By [KS98, Proposition 1.12, p. 5], there exists a $\mathbb{G}$-progressively measurable modification of the above process. $\square$

Additional assumptions We assume that the coordinator does not observe the parameters resulting from the aggregation of individual data $(\bar{\gamma}^{(k,N)}, \bar{u}^{\mathtt{ref},(k,N)}, \bar{x}^{\mathtt{ref},(k,N)}, \bar{x}_T^{\mathtt{f},(k,N)})_{k\in[M]}$, $\bar{\mathrm{P}}^{\mathtt{load},(N)}$ but can use statistical estimators of these quantities, measurable with respect to the common noise information $\mathbb{G}$. Similarly, the aggregator does not observe the empirical mean of the initial states over classes $(\frac{1}{N_k}\sum_{j=1}^{N_k} x_0^{(k,j)})_{k\in[M]}$, but can use (statistical) approximations of these parameters, denoted $(\bar{x}_0^{(k)})_{k\in[M]}$. This is realistic and desirable from the point of view of the application, as it allows to consider a non-intrusive and non omniscient coordinator.

(H-lim) Assume **(H-X)**, **(H-J)** and in addition, the following assumptions.
We denote $\pi^{(k)} := \frac{N_k}{N}$, for all $k\in[M]$ and assume $\pi^{(k)} \geq \eta$ for all $k\in[M]$ for some $\eta > 0$. In particular, $M \leq \frac{1}{\eta}$.
The meteorological conditions are represented by a stochastic process $\mathrm{X}^{\mathrm{sun}} \in \mathcal{H}$, assumed $\mathbb{G} = \{\mathcal{G}_t\}_{0\leq t\leq T}$-progressively measurable, where $\mathbb{G}$ satisfies the usual conditions and is immersed in $\mathbb{F}$.
We assume $\left|\bar{x}_0^{(k)} - \frac{1}{N_k}\sum_{j=1}^{N_k} x_0^{(k,j)}\right| \leq \frac{C}{\sqrt{N_k}} = \frac{C}{\sqrt{\pi_k N}}$, for some constant C independent from N.
$(u^{\mathtt{ref},(k,i)})_{k,i}$ (resp. $x^{\mathtt{ref},(k,i)}$, $x_T^{\mathtt{f},(k,i)}$) are independent conditionally to $\mathcal{G}_T$.
$\mathrm{P}^{\mathtt{load},(k,i)}$, $\gamma^{(k,i)}$, $u^{\mathtt{ref},(k,i)}$ and $x^{\mathtt{ref},(k,i)}$ are uniformly bounded in $\mathcal{H}$ by a constant independent from N.
Similarly, $x_T^{\mathtt{f},(k,i)}$ is bounded in $\mathbb{L}^2$ by a constant independent from N.
The processes $(\mathrm{P}^{\mathtt{load},(k,i)}, \gamma^{(k,i)}, u^{\mathtt{ref},(k,i)}, x^{\mathtt{ref},(k,i)}, x_T^{\mathtt{f},(k,i)})_{k\in[M],i\in[N_k]}$ are independent conditionally to $\mathcal{G}_T$.

Example 4.2. *Let us illustrate the assumption of independence of* $(\mathrm{P}^{\mathtt{load},(k,i)})_{k,i}$, $(u^{\mathtt{ref},(k,i)})_{k,i}$, $(x^{\mathtt{ref},(k,i)})_{k,i}$ *and* $(\gamma_{k,i})_{k,i}$ *conditionally to* $\mathbb{G}$*. For this discussion, assume* $\mathrm{P}^{\mathtt{load},(k,i)} = \mathrm{P}^{\mathrm{cons},(k,i)} - \mathrm{P}^{\mathrm{sun},(k,i)}$.

One can consider that the power consumption $\mathrm{P}^{\mathrm{cons}}$ *by a household can be explained by the temperature (as it impacts the heating and cooling), the solar irradiance (as it impacts the lighting) and other random individual factors. Our assumption states that the random individual factors are independent for distinct households. It can be interpreted as the consumers having independent behaviors, once the weather is known. Similarly, the exogenous solicitations* γ *of individual energy storage systems (thermal losses for instance), the target power and state of charge profiles* $u^{\mathtt{ref}}$ *and* $x^{\mathtt{ref}}$ *depend on meteorological conditions and individual factors. Our assumption states that these individual factors are statistically independent.*

Last, if the average solar irradiance $\mathrm{P}^{\mathrm{sun}}$ *is known in a region, there remains some random local fluctuations of solar irradiance, due to clouds passing by, for instance. Our assumption includes the case where these random local fluctuations (specific to each agent) are independent conditionally to the average solar irradiance on the region.*

4.1 Convergence of the coordination problem for large populations

When the number of agents becomes large enough, the input parameters of the *coordination problem* converge, and their limits can be easily computed, based on the conditional law of large numbers, without having to observe individual data.

Proposition 4.3. *We have the following convergence properties:*

$$\|\bar{\mathrm{P}}^{\mathtt{load},(N)} - \bar{\mathrm{P}}^{\mathtt{load}}\|_{\mathcal{H}} \leq \frac{C}{\sqrt{N}},$$
$$\forall k\in[M], \quad \|\bar{u}^{\mathtt{ref},(k,N)} - \bar{u}^{\mathtt{ref},(k)}\|_{\mathcal{H}} + \|\bar{x}^{\mathtt{ref},(k,N)} - \bar{x}^{\mathtt{ref},(k)}\|_{\mathcal{H}} + \|\bar{x}_T^{\mathtt{f},(k,N)} - \bar{x}_T^{\mathtt{f},(k)}\|_{\mathcal{H}} + \|\bar{\gamma}^{(k,N)} - \bar{\gamma}^{(k)}\|_{\mathcal{H}} \leq \frac{C}{\sqrt{N_k}}.$$

for some constant $C > 0$ independent from N.

The proof of the above Proposition is postponed to Subsection 7.3.

Replacing the coefficients of the FBSDE (3.1) by their $\mathbb{G}$-measurable approximations leads us to consider the following FBSDE with unknown $(u^{(k)}, X^{(k)}, Y^{(k)})_{k\in[M]} \in (\mathcal{H}\times\mathcal{S}\times\mathcal{S})^M$:

$$\forall k\in[M],$$
$$\begin{cases} X_t^{(k)} = \bar{x}_0^{(k)} + \int_0^t \left(\alpha_s^{(k)} u_s^{(k)} + \beta_s^{(k)} X_s^{(k)} + \bar{\gamma}_s^{(k)}\right) \mathrm{d}s, \\ Y_t^{(k)} = \mathbb{E}_t\left[\rho^{(k)}\left(X_T^{(k)} - \bar{x}_T^{\mathtt{f},(k)}\right) + \int_t^T \left(\beta_s^{(k)} Y_s^{(k)} + \nu_s^{(k)}\left(X_s^{(k)} - \bar{x}_s^{\mathtt{ref},(k)}\right)\right)\mathrm{d}s\right], \\ \mu_t^{(k)}\left(u_t^{(k)} - \bar{u}_t^{\mathtt{ref},(k)}\right) + \mathcal{L}_x'\left(t, \sum_{l=1}^M \pi^{(l)} u_t^{(l)} + \bar{\mathrm{P}}_t^{\mathtt{load}} - \mathrm{P}_t^{\mathtt{prod}}\right) + \alpha_t^{(k)} Y_t^{(k)} = 0. \end{cases} \tag{4.1}$$

This FBSDE is called the *limiting coordination problem* and has a unique solution denoted by $(\bar{u}^{(k,\infty)}, \bar{X}^{(k,\infty)}, \bar{Y}^{(k,\infty)})_{1\le k\le M} \in \mathcal{H}\times\mathcal{S}\times\mathcal{S}$. Indeed, it is the optimality system of the stochastic control problem:

$$\begin{aligned}\min_{(u^{(k)})\in\mathcal{H}^M} \quad & \mathbb{E}\left[\sum_{k=1}^{M}\pi^{(k)}\left\{\int_0^T\left(\frac{\mu_t^{(k)}}{2}\left(u_t^{(k)}-\bar{u}_t^{\mathtt{ref},(k)}\right)^2+\frac{\nu_t^{(k)}}{2}\left(X_t^{(k)}-\bar{x}_t^{\mathtt{ref},(k)}\right)^2\right)\mathrm{d}t+\frac{\rho^{(k)}}{2}\left(X_T^{(k)}-\bar{x}_T^{\mathtt{f},(k)}\right)^2\right\}\right]\\ &+\mathbb{E}\left[\int_0^T\mathcal{L}_t\left(\sum_{l=1}^{M}\pi^{(l)}u_t^{(l)}+\bar{\mathrm{P}}_t^{\mathtt{load}}-\mathrm{P}_t^{\mathtt{prod}}\right)\mathrm{d}t\right], \qquad (4.2)\\ s.t.\quad & X_t^{(k)}=\bar{x}_0^{(k)}+\int_0^t\left(\alpha_s^{(k)}u_s^{(k)}+\beta_s^{(k)}X_s^{(k)}+\bar{\gamma}_s^{(k)}\right)\mathrm{d}s,\quad \forall k\in[M].\end{aligned}$$

This control problem can be interpreted as the problem of the coordinator (which plays the role of major player in the stochastic differential game) aiming to control the aggregate behavior of the population by sending appropriate signals, in the asymptotic regime of large populations. By our assumptions and using [Bre10, Corollary 3.23, pp.71], the above control can be shown to have a unique solution, which is fully characterized by the FBSDE (4.1), by convexity of the stochastic control problem.

Definition 4.4. *Let $\left(\bar{u}^{(k,\infty)}, \bar{X}^{(k,\infty)}, \bar{Y}^{(k,\infty)}\right)\in\mathcal{H}\times\mathcal{S}\times\mathcal{S}$ be the unique solution of the limiting coordination problem* (4.1). *We define the limiting coordination signal $\bar{v}^{(\infty)}\in\mathcal{H}$ by:*

$$\forall t\in[0,T],\quad \bar{v}_t^{(\infty)}:=\mathcal{L}_x'\left(t,\sum_{l=1}^{M}\pi^{(l)}\bar{u}_t^{(l,\infty)}+\bar{\mathrm{P}}_t^{\mathtt{load}}-\mathrm{P}_t^{\mathtt{prod}}\right). \qquad (4.3)$$

The interest of considering the *limiting coordination problem* instead of the *coordination problem* is that it can be solved in the filtration $\mathbb{G}$ instead of the general filtration $\mathbb{F}$, allowing to consider a non-intrusive coordinator. This fact is made clear by the following theorem.

Theorem 4.5. *The unique solution $(\bar{u}^{(k,\infty)}, \bar{X}^{(k,\infty)}, \bar{Y}^{(k,\infty)})_{k\in[M]}\in\mathcal{H}\times\mathcal{S}\times\mathcal{S}$ of the FBSDE* (4.1) *is $\mathbb{G}$-progressively measurable.*

The proof of the above theorem is postponed to Subsection 7.4.

We then show the convergence of the solution of the *coordination problem* to the solution of *limiting coordination problem* when N goes to infinity at speed $1/\sqrt{N}$, as expected. To do it, we prove stability of the solution of FBSDE with similar structure as the one of (3.1) and (4.1) with respect to parameters of the FBSDE.

Proposition 4.6. *Let $\theta=(x,v,w,u^{\mathtt{ref}},x^{\mathtt{ref}},x_T^{\mathtt{f}})$ with $x=(x^{(k)})_{k\in[M]}\in\mathbb{R}^M$, $v\in\mathcal{H}$, $w=(w^{(k)})_{k\in[M]}\in\mathcal{H}(\mathbb{R}^M)$, $u^{\mathtt{ref}}=(u^{\mathtt{ref},(k)})_{k\in[M]}\in\mathcal{H}$, $x^{\mathtt{ref}}=(x^{\mathtt{ref},(k)})_{k\in[M]}\in\mathcal{H}$, $x_T^{\mathtt{f}}=(x_T^{\mathtt{f},(k)})_{k\in[M]}\in\mathbb{L}_T^2$.*

Consider the FBSDE with unknowns $(u^{(k)},X^{(k)},Y^{(k)})_{k\in[M]}\in\mathcal{H}\times\mathcal{S}\times\mathcal{S}$ parameterized by :

$$\begin{cases}X_t^{(k)}=x^{(k)}+\int_0^t\left(\alpha_s^{(k)}u_s^{(k)}+\beta_s^{(k)}X_s^{(k)}+w_s^{(k)}\right)\mathrm{d}s,\\ Y_t^{(k)}=\mathbb{E}_t\left[\rho^{(k)}\left(X_T^{(k)}-x_T^{\mathtt{f},(k)}\right)+\int_t^T\left(\beta_s^{(k)}Y_s^{(k)}+\nu_s^{(k)}\left(X_s^{(k)}-x_s^{\mathtt{ref},(k)}\right)\right)\mathrm{d}s\right],\\ \mu_t^{(k)}\left(u_t^{(k)}-u_t^{\mathtt{ref},(k)}\right)+\mathcal{L}_x'\left(t,\sum_{l=1}^M\pi^{(l)}u_t^{(l)}+v_t\right)+\alpha_t^{(k)}Y_t^{(k)}=0.\end{cases}$$

This FBSDE has a unique solution $(\bar{u}^{(k),\theta},\bar{X}^{(k),\theta},\bar{Y}^{(k),\theta})_{k\in[M]}\in\mathcal{H}\times\mathcal{S}\times\mathcal{S}$ for any $\theta=(x,v,w,u^{\mathtt{ref}},x^{\mathtt{ref}},x_T^{\mathtt{f}})\in\mathbb{R}^M\times\mathcal{H}\times\mathcal{H}(\mathbb{R}^M)\times\mathcal{H}(\mathbb{R}^M)\times\mathbb{L}_T^2(\mathbb{R}^M)$. Besides, for any $\theta^1=(x^1,v^1,w^1,u^{\mathtt{ref},1},x^{\mathtt{ref},1},x_T^{\mathtt{f},1})\in\mathbb{R}^M\times\mathcal{H}\times\mathcal{H}(\mathbb{R}^M)\times\mathcal{H}(\mathbb{R}^M)\times\mathbb{L}_T^2(\mathbb{R}^M)$ and $\theta^2=(x^2,v^2,w^2,u^{\mathtt{ref},2},x^{\mathtt{ref},2},x_T^{\mathtt{f},2})\in\mathbb{R}^M\times\mathcal{H}\times\mathcal{H}(\mathbb{R}^M)\times\mathcal{H}(\mathbb{R}^M)\times\mathbb{L}_T^2(\mathbb{R}^M)$, we have, for T small enough:

$$\|(\bar{u}^{\theta^1}-\bar{u}^{\theta^2},\bar{X}^{\theta^1}-\bar{X}^{\theta^2},\bar{Y}^{\theta^1}-\bar{Y}^{\theta^2})\|_{\mathcal{H}}\le C_T\|\theta^1-\theta^2\|.$$

The proof of the above Proposition is postponed to Subsection 7.5.

Remark 4.7. *This stability result could be extended to arbitrary time horizon T, for instance by an adaptation of the continuation method, presented for instance in [CD18, p. 560, Volume I], or maybe as a consequence of [Ma+15, Theorem 8.1, p. 2203].*

Corollary 4.8. *Under our assumptions, we have:*

$$\sum_{k=1}^{M} \|(\bar{u}^{(k,\infty)} - \bar{u}^{(k,N)}, \bar{X}^{(k,\infty)} - \bar{X}^{(k,N)}, \bar{Y}^{(k,\infty)} - \bar{Y}^{(k,N)})\|_{\mathcal{H}} \leq \frac{C_T}{\sqrt{N}}.$$

In particular, we have the following estimation of the error between the coordination signal and the limiting coordination signal:

$$\|\bar{v}^{(N)} - \bar{v}^{(\infty)}\|_{\mathcal{H}} \leq \frac{C_T}{\sqrt{N}}.$$

The proof of the above Corollary is postponed to Subsection 7.6.

4.2 Convergence of the individual problems for large populations

Proposition 4.9. *For $k \in [M], i \in [N_k]$, consider the following FBSDE with unknown $(u, X, Y) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$, called the limiting individual problem of agent i of class k:*

$$\begin{cases} X_t = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s + \beta_s^{(k)} X_s + \gamma_s^{(k,i)}\right) \mathrm{d}s \\ Y_t = \mathbb{E}_t \left[\rho^{(k)} \left(X_T - x_T^{\mathtt{f},(k,i)}\right) + \int_t^T \left(\beta_s^{(k)} Y_s + \nu_s^{(k)} \left(X_s - x_s^{\mathtt{ref},(k,i)}\right)\right) \mathrm{d}s\right] \\ \mu_t^{(k)} \left(u_t - u_t^{\mathtt{ref},(k,i)}\right) + \bar{v}_t^{(\infty)} + \alpha_t^{(k)} Y_t = 0 \end{cases} \tag{4.4}$$

This FBSDE has a unique solution $(u^{(k,i,\infty)}, X^{(k,i,\infty)}, Y^{(k,i,\infty)}) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$.

Proof. We use similar arguments as in the proof of Theorem 2.1. The FBSDE (4.4) is the optimality system associated to the stochastic control problem:

$$\begin{aligned} \min_{u \in \mathcal{H}} \quad & \mathbb{E}\left[\int_0^T \left(\frac{\mu_t^{(k)}}{2}\left(u_t - u_t^{\mathtt{ref},(k,i)}\right)^2 + \frac{\nu_t^{(k)}}{2}\left(X_t - x_t^{\mathtt{ref},(k,i)}\right)^2 + \bar{v}_t^{(\infty)} u_t\right) \mathrm{d}t + \frac{\rho^{(k)}}{2}\left(X_T - x_T^{\mathtt{f},(k,i)}\right)^2\right], \\ s.t. \quad & X_t = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s + \beta_s^{(k)} X_s + \gamma_s^{(k,i)}\right) \mathrm{d}s. \end{aligned} \tag{4.5}$$

By [Bre10, Corollary 3.23, pp.71], this stochastic control problem has a unique solution by our assumptions (ensuring strong convexity, continuity and coercivity of the mapping minimized in problem (4.5)). □

The stochastic control problem (4.5) can be interpreted as the individual optimization problem of agent i of class k, responding to the limiting coordination signal sent by the coordinator. Similarly as in Theorem 4.5, it could be proved that the solution of the individual problem of a given agent (k, i) is progressively measurable with respect to the completed filtration generated by the processes $x^{\mathtt{ref},(k,i)}$, $u^{\mathtt{ref},(k,i)}$, $x_T^{\mathtt{f},(k,i)}$, $\gamma^{(k,i)}$, $\bar{v}^{(\infty)}$. This is of practical interest as it ensures that individual decisions can be taken using only locally available information.

The following Proposition gives another way to compute the solution of the limiting individual problems and allows to focus on the fluctuations of the controls and states of individual players around the empirical means over classes.

Proposition 4.10. *Let $(u^{(k,i,\infty)}, X^{(k,i,\infty)}, Y^{(k,i,\infty)}) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ be the unique solution of the limiting individual problem* (4.4) *of agent $i \in [N_k]$ of class $k \in [M]$. Let $(\bar{u}^{(k,\infty)}, \bar{X}^{(k,\infty)}, \bar{Y}^{(k,\infty)})_{k \in [M]} \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ be the unique solution of the limiting coordination problem* (4.1). *Then $(u^{(k,i,\infty)} - \bar{u}^{(k,\infty)}, X^{(k,i,\infty)} - \bar{X}^{(k,\infty)}, Y^{(k,i,\infty)} - \bar{Y}^{(k,\infty)}) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ is the unique solution of the following FBSDE with unknowns $(\Delta u, \Delta X, \Delta Y) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$:*

$$\forall t \in [0, T],$$
$$\begin{cases} \Delta X_t = x_0^{(k,i)} - \bar{x}_0^{(k)} + \int_0^t \left(\alpha_s^{(k)} \Delta u_s + \beta_s^{(k)} \Delta X_s + \gamma_s^{(k,i)} - \bar{\gamma}_s^{(k)}\right) \mathrm{d}s, \\ \Delta Y_t = \mathbb{E}_t \left[\rho^{(k)} \left(\Delta X_T - x_T^{\mathtt{f},(k,i)} + \bar{x}_T^{\mathtt{f},(k)}\right) + \int_t^T \left(\beta_s^{(k)} \Delta Y_s + \nu_s^{(k)} \left(\Delta X_s - x_s^{\mathtt{ref},(k,i)} + \bar{x}_s^{\mathtt{ref},(k)}\right)\right) \mathrm{d}s\right], \\ \mu_t^{(k)} \left(\Delta u_t - u_t^{\mathtt{ref},(k,i)} + \bar{u}_t^{\mathtt{ref},(k)}\right) + \alpha_t^{(k)} \Delta Y_t = 0. \end{cases} \tag{4.6}$$

To show the convergence of the solutions of the *individual problems* to the solutions of the *limiting individual problems*, we have to use the following stability result for a class of FBSDE with respect to its input parameters.

Proposition 4.11. *Let $v \in \mathcal{H}$. For $k \in [M], i \in [N_k]$, consider the FBSDE with unknowns $(u, X, Y) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$:*

$$\begin{cases} X_t = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s + \beta_s^{(k)} X_s + \gamma_s^{(k,i)}\right) \mathrm{d}s, \\ Y_t = \mathbb{E}_t \left[\rho^{(k)} \left(X_T - x_T^{\mathtt{f},(k,i)}\right) + \int_t^T \left(\beta_s^{(k)} Y_s + \nu_s^{(k)} \left(X_s - x_s^{\mathtt{ref},(k,i)}\right)\right) \mathrm{d}s\right], \\ \mu_t^{(k)} \left(u_t - u_t^{\mathtt{ref},(k,i)}\right) + v_t + \alpha_t^{(k)} Y_t = 0. \end{cases}$$

This FBSDE has a unique solution $(u^{(k,i),v}, X^{(k,i),v}, Y^{(k,i),v}) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$ for any $v \in \mathcal{H}$. Besides, for T small enough, for any v, v' in $\mathcal{H}$, for all $k \in [M], i \in [N_k]$:

$$\|(u^{(k,i),v} - u^{(k,i),v'}, X^{(k,i),v} - X^{(k,i),v'}, Y^{(k,i),v} - Y^{(k,i),v'})\|_{\mathcal{H}} \leq C_T \|v - v'\|_{\mathcal{H}}.$$

The proof of the above Proposition is similar to the proof of Proposition 4.6.

Theorem 4.12. *The solutions of the limiting individual problems are close to the solutions of the individual problems. In other words, for $k \in [M], i \in [N_k]$, for T small enough, for some constant C independent from N,*

$$\|(u^{(k,i,\infty)} - u^{(k,i,N)}, X^{(k,i,\infty)} - X^{(k,i,N)}, Y^{(k,i,\infty)} - Y^{(k,i,N)})\|_{\mathcal{H}} \leq \frac{C}{\sqrt{N}},$$

and we have the following estimation on the sub-optimality of $u^{(\infty)} := (u^{(k,i,\infty)})_{k \in [M], i \in [N_k]}$ compared to the optimal solution $u^{(N)} := (u^{(k,i,N)})_{k \in [M], i \in [N_k]}$, for another constant C independent from N:

$$0 \leq \mathcal{J}(u^{(\infty)}) - \mathcal{J}(u^{(N)}) = \mathcal{J}(u^{(\infty)}) - \min_{v \in \mathcal{H}} \mathcal{J}(v) \leq \frac{C}{N}.$$

The proof of the above Theorem is postponed to Subsection 7.7.

The above result shows that the (unique) Nash equilibrium a Stochastic Stackelberg Differential Game with followers (the agents) and a leader (the aggregator) corresponds to a quasi-optimal solution of the centralized control problem of many TCLs. In particular, no real-time communication from the agents to the aggregator is required, and the problems of the agents and the coordinator can be solved using locally available data only.

5 Numerical experiments

5.1 The model

5.1.1 Models for exogenous random uncertainties

In the following, $M = 2$, $W = (W^{(k,i)})_{k \in [2], i \in [N_k]}$, $\tilde{W}$ and $N^{\text{cons}} = (N_{k,i}^{\text{cons},(k,i)})_{k \in [2], i \in [N_k]}$ denote respectively a N-dimensional Brownian motion, a one-dimensional Brownian motion and a N independent compensated Poisson processes with intensity λ^{cons}. These processes are assumed independent.

We assume $\mathtt{P}^{\mathtt{load},(k,i)} = \mathtt{P}^{\text{cons},(k,i)} - \mathtt{P}^{\text{sun}}$ for all $k \in [2], i \in N_k$. For the consumption process $\mathtt{P}^{\text{cons},(k,i)}$, we assume the following dynamic:

$$\mathrm{d}\mathtt{P}_t^{\text{cons},(k,i)} = -\rho^{\text{cons}} (\mathtt{P}_t^{\text{cons},(k,i)} - \mathtt{p}_t^{\text{cons},\mathtt{ref}}) \mathrm{d}t + \sigma_t^{\text{cons}} \mathrm{d}W_t^{(k,i)} + h^{\text{cons}} \mathrm{d}N_t^{\text{cons},(k,i)}. \tag{5.1}$$

In practice, only conditional independence of the processes $(\mathtt{P}^{\text{cons},(k,i)})$ is necessary for our results to hold, but for simplicity of our presentation, we assume that they are (unconditionally) independent identically distributed.

Regarding the PV production, we follow [Bad+18] by setting $\mathtt{P}^{\text{sun}} = \mathtt{P}^{\text{sun,max}} \mathtt{X}^{\text{sun}}$ where $\mathtt{P}^{\text{sun,max}} : [0, T] \mapsto \mathbb{R}$ is a deterministic function (the clear sky model) and $\mathtt{X}^{\text{sun}}$ solves a Fisher-Wright type SDE which dynamics is

$$\mathrm{d}\mathtt{X}_t^{\text{sun}} = -\rho^{\text{sun}} (\mathtt{X}_t^{\text{sun}} - \mathtt{x}_t^{\text{sun},\mathtt{ref}}) \mathrm{d}t + \sigma^{\text{sun}} (\mathtt{X}_t^{\text{sun}})^{k_1} (1 - \mathtt{X}_t^{\text{sun}})^{k_2} \mathrm{d}\tilde{W}_t, \tag{5.2}$$

with $k_1, k_2 \geq 1/2$. As proved in [Bad+18], there is a strong solution to the above SDE and the solution X^{sun} takes values in $[0, 1]$.

Since the drifts are affine-linear, the conditional expectations of the solutions are known in closed form (this property is intensively used in [BSS05]):

$$\mathbb{E}_t\left[P_s^{\text{sun}}\right] = \left(\frac{P_t^{\text{sun}}}{P_t^{\text{sun,max}}}\exp(-\rho^{\text{sun}}(s-t)) + \int_t^s \rho^{\text{sun}} x_\tau^{\text{sun,ref}} \exp(-\rho^{\text{sun}}(s-\tau))\mathrm{d}\tau\right) P_s^{\text{sun,max}}, \tag{5.3}$$

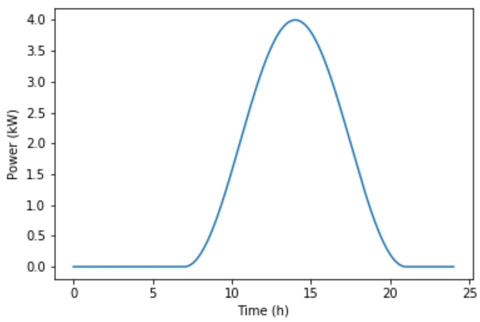
$$\mathbb{E}_t\left[P_s^{\text{cons},(k,i)}\right] = P_t^{\text{cons},(k,i)}\exp(-\rho^{\text{cons}}(s-t)) + \left(\int_t^s \rho^{\text{cons}} p_\tau^{\text{cons,ref}} \exp(-\rho^{\text{cons}}(s-\tau))\mathrm{d}\tau\right), \tag{5.4}$$

for $s \geq t$. This will allow us to speed up computations of the conditional expectations $\mathbb{E}_t\left[\bar{P}_s^{\text{load}}\right]$ and $\mathbb{E}_t\left[\bar{P}_s^{\text{load},(N)}\right]$ as required when deriving the optimal control. Moreover, throughout our application, we assume $P_t^{\text{prod}} = \mathbb{E}\left[\bar{P}_t^{\text{load},(N)}\right]$, which can be easily computed using the previous remark.

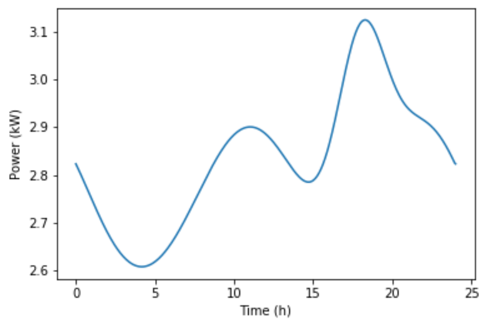
The values of the parameters used are given in the following table, while $P^{\text{sun,max}}$ and $p^{\text{cons,ref}}$ are plotted in Figures 2a and 2b. Empirical quantile plot (obtained by simulation of 40000 i.i.d. trajectories) as well as one example trajectory of P^{sun} and P^{cons} are given in Figures 2c and 2d. The simulations are performed using Euler scheme on a time horizon $T = 24$ h with step length $1/16$ h.

Table 1: Parameter values for the simulation of input processes

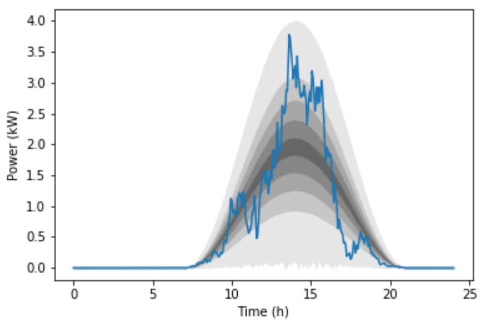
ρ^{cons}	σ^{cons}	λ^{cons}	h^{cons}	ρ^{sun}	$x^{\text{sun,ref}}$	σ^{sun}	k_1	k_2
0.9 h^{-1}	$0.1(p^{\text{cons,ref}} - 2.5) + 0.1$	0.25 h^{-1}	0.25 kW	0.75 h^{-1}	0.5	0.8	0.8	0.7



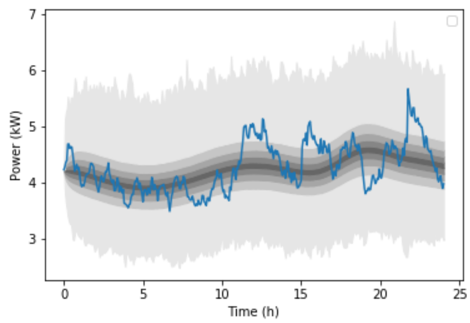
(a) Time evolution of $P^{\text{sun,max}}$



(b) Time evolution of $p^{\text{cons,ref}}$



(c) Quantile plot and one example trajectory of P^{sun}



(d) Quantile plot and one example trajectory of P^{cons}

Figure 2: Parameters and statistical evolution of P^{sun} and P^{cons}

5.1.2 Models for Thermostatically Controlled Loads

We consider two classes (i.e., $M = 2$) of Thermostatically Controlled Loads: heat pumps and water heaters. We control the power consumption of these devices around a nominal value defined such that thermal losses are exactly compensated when their target temperature $x_T^{\mathtt{f},(k,i)} = x^{\mathtt{ref},(k,i)}$, assumed constant, is reached. We assume first order models for the temperature, as in [Tro+16]. The dynamics of the temperature associated to individual devices is given by:

$$\frac{\mathrm{d}X_t^{(k,i)}}{\mathrm{d}t} = \frac{COP^{(k)}}{C^{(k)}} u_t^{(k,i)} - \frac{1}{R^{(k)}C^{(k)}}(X_t^{(k,i)} - x^{\mathtt{ref},(k,i)}),$$

where $COP^{(k)}$ denotes the coefficient of performance of devices of type k, $C^{(k)}$ denotes its thermal capacitance, $R^{(k)}$ its thermal resistance, $X^{out,(k,i)}$ the environment temperature of the device. This gives $\alpha^{(k)} = \frac{COP^{(k)}}{C^{(k)}}$, $\beta^{(k)} = -\frac{1}{R^{(k)}C^{(k)}}$ and $\gamma^{(k,i)} = \frac{x^{\mathtt{ref},(k,i)}}{R^{(k)}C^{(k)}}$. Realistic parameter values for various types of TCLs (AC, refrigerators, water heaters, heat pumps) can be found in [Mat+12]. We use these values to set the parameters of our models of devices.

Table 2: Parameter values for the Thermostatically Control Loads

Class index k	type of device	$R^{(k)}$	$C^{(k)}$	$COP^{(k)}$	$\bar{x}^{\mathtt{ref},(k)}$	Deadband $\delta^{(k)}$
1	Water heater	120 (°C/kW)	0.4 (kWh/ °C)	1	48 (°C)	4 (°C)
2	Heat pump	2 (°C/kW)	2 (kWh/°C)	3.5	20 (°C)	1 (°C)

The temperature deadband $\delta^{(k)}$ of a TCLs denotes admissible deviation of the temperature $X^{(k,i)}$ of a device from its reference $x^{\mathtt{ref},(k,i)}$. We assume $(x_0^{(k,i)})_{k\in[2],i\in[N_k]} = (x^{\mathtt{ref},(k,i)})_{k\in[2],i\in[N_k]} = (x_T^{\mathtt{f},(k,i)})_{k\in[2],i\in[N_k]}$ are independent and $x^{\mathtt{ref},(k,i)} = x_0^{(k,i)} = x_T^{\mathtt{f},(k,i)}$ is drawn from a uniform distribution on the interval $[\bar{x}^{\mathtt{ref},(k)} - \delta^{(k)}, \bar{x}^{\mathtt{ref},(k)} + \delta^{(k)}]$.

5.1.3 Cost parameters

In the case where $\mathcal{L}$ is not quadratic, the coordination problem is a low-dimensional control problem which does not have explicit solutions in general, so that a numerical method is required to solve it. We assume in this paper that $\mathcal{L}$ is simply the quadratic function $\mathcal{L} : (t, x) \mapsto \frac{\lambda}{2}x^2$ for some deterministic constant $\lambda \geq 0$. In this case, all the control problems are Linear-quadratic, the associated FBSDEs are affine-linear, and therefore have quasi-explicit solutions. Individual cost parameters are scaled according to the square of the temperature deadband of the type of device considered, in order to guarantee proper scaling of the temperature range sizes.

Table 3: Parameter values for the cost functional

$\mu^{(k)}(\delta^{(k)})^2$	$\nu^{(k)}(\delta^{(k)})^2$	$\rho^{(k)}(\delta^{(k)})^2$	λ
4 (°C)^2 kW^{-2} h^{-1}	5 h^{-1}	0	30 kW^{-2} h^{-1}

5.2 Solution method of affine-linear FBSDEs

The solution method for linear FBSDEs is standard, see for instance [Bis76; Yon99; Yon06].

5.2.1 The limiting coordination problem

The optimality system of the limiting coordination problem is a linear FBSDE which writes, up to rescaling of the optimality condition by $\boldsymbol{\pi}$:

$$\begin{cases} X_0 = \bar{x}_0, \\ \frac{\mathrm{d}X_t}{\mathrm{d}t} = \boldsymbol{\alpha} u_t + \boldsymbol{\beta} X_t + \bar{\gamma}, \\ Y_t = \mathbb{E}_t\left[\boldsymbol{\rho}(X_T - \bar{x}^{\mathrm{ref}}) + \int_t^T \left(\boldsymbol{\nu}(X_s - \bar{x}^{\mathrm{ref}}) + \boldsymbol{\beta} Y_s\right)\mathrm{d}s\right], \\ \boldsymbol{\mu}\boldsymbol{\pi} u_t + \lambda\pi(\pi^\top u_t + \bar{\mathrm{P}}_t^{\mathrm{load}} - \mathrm{P}_t^{\mathrm{prod}}) + \boldsymbol{\alpha}\boldsymbol{\pi} Y_t = 0, \end{cases} \tag{5.5}$$

where:

$$\boldsymbol{\alpha} := diag(\alpha^{(k)})_{k=1,2}, \quad \boldsymbol{\beta} := diag(\beta^{(k)})_{k=1,2}, \quad \boldsymbol{\mu} := diag(\mu^{(k)})_{k=1,2}, \quad \boldsymbol{\nu} := diag(\nu^{(k)})_{k=1,2}, \quad \boldsymbol{\rho} := diag(\rho^{(k)})_{k=1,2},$$

$$\pi := \begin{pmatrix} \pi^{(1)} \\ \pi^{(2)} \end{pmatrix}, \quad \boldsymbol{\pi} := diag(\pi), \quad \bar{\gamma} = (\bar{\gamma}^{(k)})_{k=1,2}, \quad \bar{u} = (\bar{u}^{(k)})_{k=1,2}, \quad \bar{X} = (\bar{X}^{(k)})_{k=1,2}, \quad \bar{Y} = (\bar{Y}^{(k)})_{k=1,2}.$$

The last equation allows to eliminate u, which writes:

$$u_t = -(\boldsymbol{\mu}\boldsymbol{\pi} + \lambda\pi\pi^\top)^{-1}\left(\boldsymbol{\alpha}\boldsymbol{\pi} Y_t + \lambda(\bar{\mathrm{P}}_t^{\mathrm{load}} - \mathrm{P}_t^{\mathrm{prod}})\pi\right).$$

Therefore, we obtain the following FBSDE with unknowns (X, Y):

$$\begin{cases} X_0 = \bar{x}_0, \\ \frac{\mathrm{d}X_t}{\mathrm{d}t} = \boldsymbol{\beta} X_t - \boldsymbol{\alpha}(\boldsymbol{\mu}\boldsymbol{\pi} + \lambda\pi\pi^\top)^{-1}\boldsymbol{\alpha}\boldsymbol{\pi} Y_t + \bar{\gamma} + \lambda\boldsymbol{\alpha}(\boldsymbol{\mu}\boldsymbol{\pi} + \lambda\pi\pi^\top)^{-1}(\mathrm{P}_t^{\mathrm{prod}} - \bar{\mathrm{P}}_t^{\mathrm{load}})\pi, \\ Y_t = \mathbb{E}_t\left[\boldsymbol{\rho}(X_T - \bar{x}^{\mathrm{ref}}) + \int_t^T \left(\boldsymbol{\nu}(X_s - \bar{x}^{\mathrm{ref}}) + \boldsymbol{\beta} Y_s\right)\mathrm{d}s\right]. \end{cases} \tag{5.6}$$

Introduce the Matrix-valued Riccati ODE with unknown ϕ:

$$\begin{cases} \frac{\mathrm{d}\phi_t}{\mathrm{d}t} + \phi_t\boldsymbol{\beta} + \boldsymbol{\beta}\phi_t + \boldsymbol{\nu}\boldsymbol{\pi} - \phi_t\boldsymbol{\alpha}(\boldsymbol{\mu}\boldsymbol{\pi} + \lambda\pi\pi^\top)^{-1}\boldsymbol{\alpha}\phi_t = 0, \\ \phi_T = \boldsymbol{\rho}\boldsymbol{\pi}. \end{cases}$$

According to [Bis76, Theorem 6.1], this Riccati ODE has a unique bounded solution $\bar{\boldsymbol{\phi}}$, since $\boldsymbol{\rho}\boldsymbol{\pi}$ and $\boldsymbol{\mu}\boldsymbol{\pi} + \lambda\pi\pi^\top$ are positive semi-definite matrices (the second matrix is even positive definite).

Introduce the linear BSDE in $(\psi, M) \in \mathcal{S} \times \mathcal{S}$ with M martingale vanishing at $t = 0$:

$$\begin{cases} -\mathrm{d}\psi_t = \left\{\left(\boldsymbol{\beta} - \bar{\boldsymbol{\phi}}_t\boldsymbol{\alpha}(\boldsymbol{\mu}\boldsymbol{\pi} + \lambda\pi\pi^\top)^{-1}\boldsymbol{\alpha}\right)\psi_t - \boldsymbol{\nu}\boldsymbol{\pi}\bar{x}^{\mathrm{ref}} + \bar{\boldsymbol{\phi}}_t\bar{\gamma} + \lambda\bar{\boldsymbol{\phi}}_t\boldsymbol{\alpha}(\boldsymbol{\mu}\boldsymbol{\pi} + \lambda\pi\pi^\top)^{-1}(\mathrm{P}_t^{\mathrm{prod}} - \bar{\mathrm{P}}_t^{\mathrm{load}})\pi\right\}\mathrm{d}t - \mathrm{d}M_t, \\ \psi_T = -\boldsymbol{\rho}\boldsymbol{\pi}\bar{x}^{\mathrm{ref}}. \end{cases}$$

This linear BSDE admits a unique solution $(\bar{\psi}, \bar{M}) \in \mathcal{S} \times \mathcal{S}$ with $\bar{M}$ martingale vanishing at $t = 0$, according to [EPQ97, Theorem 5.1]. Let $\bar{X}$ be the unique solution of the affine-linear ODE:

$$\begin{cases} X_0 = \bar{x}_0, \\ \frac{\mathrm{d}X_t}{\mathrm{d}t} = \boldsymbol{\beta} X_t - \boldsymbol{\alpha}(\boldsymbol{\mu}\boldsymbol{\pi} + \lambda\pi\pi^\top)^{-1}\boldsymbol{\alpha}(\bar{\boldsymbol{\phi}}_t X_t + \bar{\psi}_t) + \bar{\gamma} + \lambda\boldsymbol{\alpha}(\boldsymbol{\mu}\boldsymbol{\pi} + \lambda\pi\pi^\top)^{-1}(\mathrm{P}_t^{\mathrm{prod}} - \bar{\mathrm{P}}_t^{\mathrm{load}})\pi. \end{cases}$$

Then, using Integration by Parts Formula in [Pro03, Corollary 2, p. 68], one can show that $\bar{X}$ and $\bar{Y} = \boldsymbol{\pi}^{-1}\left(\bar{\boldsymbol{\phi}}\bar{X} + \bar{\psi}\right)$ are solutions of the FBSDE (5.6). Therefore, (5.5) has a unique solution $(\bar{u}, \bar{X}, \bar{Y})$ with the following feedback expression for $\bar{u}$:

$$\bar{u}_t = -(\boldsymbol{\mu}\boldsymbol{\pi} + \lambda\pi\pi^\top)^{-1}\boldsymbol{\alpha}(\bar{\boldsymbol{\phi}}_t\bar{X}_t + \bar{\psi}_t) + \lambda(\boldsymbol{\mu}\boldsymbol{\pi} + \lambda\pi\pi^\top)^{-1}\pi(\mathrm{P}_t^{\mathrm{prod}} - \bar{\mathrm{P}}_t^{\mathrm{load}}).$$

Besides, the limiting coordination signal is given by the feedback expression:

$$\begin{aligned} \bar{v}_t^{(\infty)} &= \lambda\left(\pi^\top\bar{u}_t + \bar{\mathrm{P}}_t^{\mathrm{load}} - \mathrm{P}_t^{\mathrm{prod}}\right) \\ &= \lambda\left(-\pi^\top(\boldsymbol{\mu}\boldsymbol{\pi} + \lambda\pi\pi^\top)^{-1}\boldsymbol{\alpha}(\bar{\boldsymbol{\phi}}_t\bar{X}_t + \bar{\psi}_t) + \lambda\pi^\top(\boldsymbol{\mu}\boldsymbol{\pi} + \lambda\pi\pi^\top)^{-1}\pi(\mathrm{P}_t^{\mathrm{prod}} - \bar{\mathrm{P}}_t^{\mathrm{load}}) + \bar{\mathrm{P}}_t^{\mathrm{load}} - \mathrm{P}_t^{\mathrm{prod}}\right). \end{aligned}$$

5.2.2 The coordination problem

The coordination problem has the same structure as the limiting coordination problem with coefficients $\bar{x}^{\mathtt{ref}}$, $\bar{x}_0$, $\bar{\gamma}$ and $\bar{\mathrm{P}}^{\mathtt{load}}$ formally replaced respectively by $\bar{x}^{\mathtt{ref},(N)}$, $\bar{x}_0^{(N)}$, $\bar{\gamma}^{(N)}$ and $\bar{\mathrm{P}}^{\mathtt{load},(N)}$. Solving the coordination problem can be done similarly as for the limiting coordination problem, by formally replacing parameters $\bar{x}^{\mathtt{ref}}$, $\bar{x}_0$, $\bar{\gamma}$ and $\bar{\mathrm{P}}^{\mathtt{load}}$ by $\bar{x}^{\mathtt{ref},(N)}$, $\bar{x}_0^{(N)}$, $\bar{\gamma}^{(N)}$ and $\bar{\mathrm{P}}^{\mathtt{load},(N)}$.

5.2.3 The individual and limiting individual problems

Once the solution of the coordination problem is computed, solving the individual and limiting individual problems amounts to solve (3.5) and (4.6). By eliminating the (recentered) control using the last equation in both these FBSDEs, one can show that the recentered individual problems and recentered limiting individual problems are both equivalent to one-dimensional FBSDEs with the following structure:

$$\begin{cases} X_t = x_0 + \int_0^t (AX_s + BY_s + a_s)\mathrm{d}s, \\ Y_t = \mathbb{E}_t\left[\Gamma X_T + f + \int_t^T (CX_s + AY_s + b_s)\mathrm{d}s\right], \end{cases} \tag{5.7}$$

with A, B, C, Γ deterministic, $B < 0$, $C, \Gamma \geq 0$, a, b in $\mathcal{H}$, f in $\mathbb{L}^2_{\mathcal{F}_T}(\Omega)$. The structure of the online limiting individual problems is similar.

Lemma 5.1 (1-dimension Riccati ODE with constant coefficients)**.** *Consider the following Riccati ODE:*

$$\begin{cases} \frac{\mathrm{d}\phi_t}{\mathrm{d}t} + a\phi_t + b\phi_t^2 + c = 0, \\ \phi_T = \gamma, \end{cases} \tag{5.8}$$

with a, b, c, γ deterministic, $b < 0$, $c, \gamma \geq 0$. Then this equation admits a unique bounded solution on $[0, T]$, denoted by ϕ. Define θ as the unique solution of the second-order linear ODE:

$$\begin{cases} \frac{\mathrm{d}^2\theta_t}{\mathrm{d}t^2} + a\frac{\mathrm{d}\theta_t}{\mathrm{d}t} + bc\theta_t = 0, \\ \frac{\mathrm{d}\theta_T}{\mathrm{d}t} = \gamma b, \\ \theta_T = 1. \end{cases} \tag{5.9}$$

Then θ is positive and the unique solution ϕ of (5.8) *on $[0, T]$ is given by:*

$$\forall t \in [0, T], \quad \phi_t = \left(\frac{\mathrm{d}\theta_t}{\mathrm{d}t}\right)\frac{1}{b\theta_t}$$

Proof. As $x \mapsto -ax - bx^2 - c$ is locally Lipschitz-continuous, (5.8) has a unique solution ϕ on some maximal interval (t_0, t_1) with $t_0 < T < t_1$ and $t_0 \in \bar{\mathbb{R}}$ and $t_1 \in \bar{\mathbb{R}}$ are unique. Consider the following Riccati ODE:

$$\begin{cases} \frac{\mathrm{d}p_t}{\mathrm{d}t} + ap_t + bp_t^2 = 0, \\ p_T = 0. \end{cases} \tag{5.10}$$

The null function is the unique solution of 5.10 on $(-\infty, +\infty)$. Consider as well the following linear ODE:

$$\begin{cases} \frac{\mathrm{d}p_t}{\mathrm{d}t} + ap_t + c = 0, \\ p_T = \gamma. \end{cases}$$

It admits a unique solution $\overline{\phi}$ on $(-\infty, +\infty)$. Besides, by comparison theorem for Ordinary differential equations, we have :

$$\forall t \in (t_0, t_1), \quad 0 \leq \phi_t \leq \overline{\phi}_t$$

which shows that ϕ can not explode in finite time, and hence $t_0 = -\infty$ and $t_1 = +\infty$. Hence ϕ is well-defined and bounded on $[0, T]$. Now, let us define θ as the unique solution of the following ODE:

$$\begin{cases} \frac{\mathrm{d}\theta_t}{\mathrm{d}t} = b\phi_t\theta_t, \\ \theta_T = 1. \end{cases} \tag{5.11}$$

Then we immediately get that θ is well-defined, and positive on $\mathbb{R}$. Besides, θ is C^2 and:

$$\begin{aligned} \frac{\mathrm{d}^2\theta_t}{\mathrm{d}t^2} &= b\frac{\mathrm{d}\phi_t}{\mathrm{d}t}\theta_t + b\phi_t\frac{\mathrm{d}\theta_t}{\mathrm{d}t} \\ &= b(-a\phi_t - b\phi_t^2 - c)\theta_t + b^2\phi_t^2\theta_t \\ &= -ab\phi_t\theta_t - bc\theta_t \\ &= -a\frac{\mathrm{d}\theta_t}{\mathrm{d}t} - bc\theta_t, \end{aligned}$$

where we used successively that ϕ solves (5.8) and that θ solves (5.11). In particular, this shows that θ is also the unique solution of (5.9). This completes the proof. □

Theorem 5.2 (Verification theorem for affine-linear FBSDE with constant coefficients)**.** *Let A, B, C, Γ be deterministic constants, $B < 0$, $C, \Gamma \geq 0$, a, b in $\mathcal{H}$, f in $\mathbb{L}^2_{\mathcal{F}_T}(\Omega)$. Let ϕ be the unique solution of the following Riccati ODE:*

$$\begin{cases} \frac{\mathrm{d}\phi_t}{\mathrm{d}t} + 2A\phi_t + B\phi_t^2 + C = 0, \\ \phi_T = \Gamma, \end{cases} \tag{5.12}$$

and let $(\psi, M) \in \mathcal{S} \times \mathcal{M}_0^2$ be the unique solution of the following BSDE:

$$\begin{cases} -\mathrm{d}\psi_t = \Big((B\phi_t + A)\psi_t + \phi_t a_t + b_t\Big)\mathrm{d}t - \mathrm{d}M_t, \\ \psi_T = f, \end{cases} \tag{5.13}$$

where $\mathcal{M}_0^2$ denotes the space of martingales in $\mathcal{S}$ vanishing at $t = 0$. Denoting θ the unique (non-negative) solution of:

$$\begin{cases} \frac{\mathrm{d}^2\theta_t}{\mathrm{d}t^2} + 2A\frac{\mathrm{d}\theta_t}{\mathrm{d}t} + BC\theta_t = 0, \\ \frac{\mathrm{d}\theta_T}{\mathrm{d}t} = \Gamma B, \\ \theta_T = 1, \end{cases} \tag{5.14}$$

we have the explicit formula for ψ:

$$\psi_t = \mathbb{E}_t\,[f]\Big(\frac{\theta_T}{\theta_t}\Big)\exp(A(T-t)) + \mathbb{E}_t\bigg[\int_t^T (a_s\phi_s + b_s)\Big(\frac{\theta_s}{\theta_t}\Big)\exp(A(s-t))\mathrm{d}s\bigg].$$

If θ is the unique solution of (5.14)*, define X by:*

$$X_t = x_0\frac{\theta_t}{\theta_0}\exp(At) + \int_0^t (B\psi_s + a_s)\frac{\theta_t}{\theta_s}\exp(A(t-s))\mathrm{d}s. \tag{5.15}$$

Define also $Y := \phi X + \psi$. Then $(X, Y) \in \mathcal{S} \times \mathcal{S}$ is a solution of the following FBSDE:

$$\begin{cases} X_t = x_0 + \int_0^t (AX_s + BY_s + a_s)\mathrm{d}s, \\ Y_t = \mathbb{E}_t\Big[\Gamma X_T + f + \int_t^T (CX_s + AY_s + b_s)\mathrm{d}s\Big]. \end{cases} \tag{5.16}$$

Proof. By Lemma 5.1, the Riccati ordinary differential equation (5.12) has a unique solution ϕ. The uniqueness of the solution $(\psi, M) \in \mathcal{S} \times \mathcal{M}_0$ of (5.13) arises from an application of [EPQ97, Theorem 5.1, p. 54]. To obtain the explicit expression of ψ, we use the Integration by Parts Formula in [Pro03, Corollary 2, p. 68] to the product

$\tilde{\psi}$ defined by $\tilde{\psi}_t := \psi_t \exp\left(-\int_t^T (B\phi_s + A)\mathrm{d}s\right)$ between 0 and T, using the fact that the second term is bounded, continuous with finite variations. This shows:

$$\begin{cases} -\mathrm{d}\tilde{\psi}_t = (\phi_t a_t + b_t)\exp\left(-\int_t^T (B\phi_s + A)\mathrm{d}s\right)\mathrm{d}t - \exp\left(-\int_t^T (B\phi_s + A)\mathrm{d}s\right)\mathrm{d}M_t, \\ \tilde{\psi}_T = f. \end{cases}$$

In particular, using the boundedness of ϕ, the last term in the above BSDE is a true martingale, so that:

$$\tilde{\psi}_t = \mathbb{E}_t\left[f + \int_t^T (\phi_s a_s + b_s)\exp\left(-\int_s^T (B\phi_r + A)\mathrm{d}r\right)\mathrm{d}s\right].$$

We obtain the explicit expression of ψ by using $\psi_t = \tilde{\psi}_t \exp\left(\int_t^T (B\phi_s + A)\mathrm{d}s\right)$ and using the fact that $\phi_t = \frac{\mathrm{d}\theta_t}{\mathrm{d}t}\frac{1}{B\theta_t}$, which yields $\exp\left(\int_t^s (B\phi_r + A)\mathrm{d}r\right) = \frac{\theta_s}{\theta_t}\exp(A(s-t))$, since θ is positive. Let $\hat{X}$ be given by:

$$\hat{X}_t = x_0 + \int_0^t ((A + B\phi_s)\hat{X}_s + B\psi_s + a_s)\mathrm{d}s. \tag{5.17}$$

We want to show that $\hat{X} = X$ given in (5.15). Define $\tilde{X}$ by $\tilde{X}_t = \hat{X}_t \exp\left(-\int_0^t (\phi_s B + A)\mathrm{d}s\right)$. Then, by integration by part, we obtain $\tilde{X}_t = x_0 + \int_0^t (B\psi_s + a_s)\exp\left(-\int_0^s (\phi_r B + A)\mathrm{d}r\right)\mathrm{d}s$. We can finally show that $\hat{X} = X$ using $\hat{X}_t = \tilde{X}_t \exp\left(\int_0^t (\phi_s B + A)\mathrm{d}s\right)$ and $\exp\left(\int_s^t (B\phi_r + A)\mathrm{d}r\right) = \frac{\theta_t}{\theta_s}\exp(A(t-s))$ for any t and s in $[0, T]$, using the form $\phi = \frac{\dot{\theta}}{B\theta}$ given by Lemma 5.1 (where $\dot{\theta}$ is the derivative of $t \mapsto \theta_t$). Then using the definition of $Y := \phi X + \psi$ and by (5.17), we get that X satisfies:

$$X_t = x_0 + \int_0^t (AX_s + BY_s + a_s)\mathrm{d}s.$$

One can verify using $Y = \phi X + \psi$ and an integration by parts that $(Y, M) \in \mathcal{S} \times \mathcal{M}_0^2$ is solution of the following BSDE:

$$\begin{cases} -\mathrm{d}Y_t = (CX_t + AY_t + b_t)\mathrm{d}t - \mathrm{d}M_t, \\ Y_T = \Gamma X_T + f, \end{cases}$$

which yields the result. □

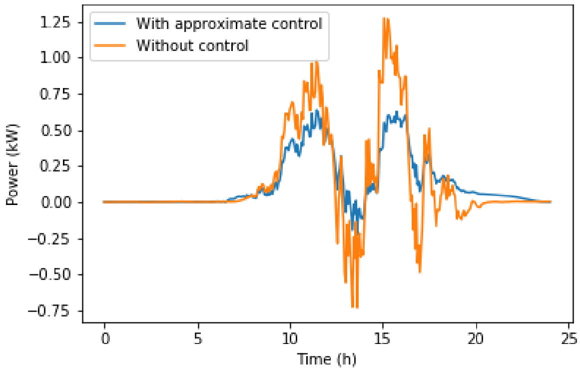
In the simulations, we rely heavily on Theorem 5.2 to solve the one-dimensional FBSDEs.

5.3 Numerical simulations and results

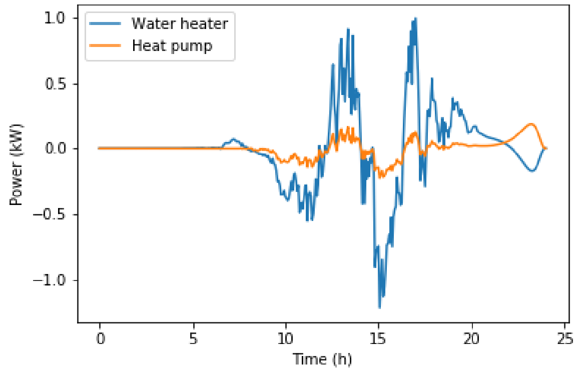
The simulations have been performed on Python 3.7, with an Intel-Core i7 PC at 2.1 GHz with 16 Go memory. We simulate one realization of the stochastic process $\mathrm{P}^{\mathrm{sun}}$ and N i.i.d scenarios of $\mathrm{P}^{\mathrm{cons}}$ (N being the number of agents) on $[0, T]$ with $T = 24$ hours using Euler schemes with step length $1/16$ h. The solutions of ordinary differential equation and linear backward stochastic differential equations are computed using Euler scheme. For the solution of linear BSDE, we rely heavily on the assumption of affine-linear processes given in Section 5.1.1.

5.3.1 Results with identical population sizes

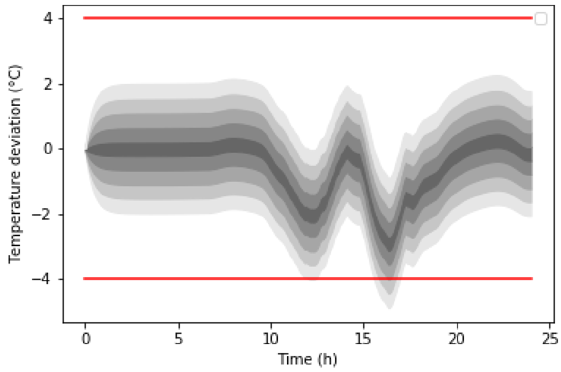
We consider $N = 40000$ users, with $N_1 = N_2 = 20000$ users in each class (which yields the relative population sizes $\pi_1 = \pi_2 = 0.5$). The results of the simulation for one weather scenario are given in Figure 3. In particular, the first graph 3a shows that the power imbalance using the approximate control (obtained by solving the limiting coordination and limiting individual problems) is closer to 0 than without control. This is done without violating temperature bounds for the populations of water heaters and heat pumps (at least not often and with low probability), see Figures 5c and 5d. This shows the interest of our approach: power imbalance may be reduced by distributed TCLs while guaranteeing good quality of service, i.e., while maintaining the temperatures of the devices in their admissible ranges.



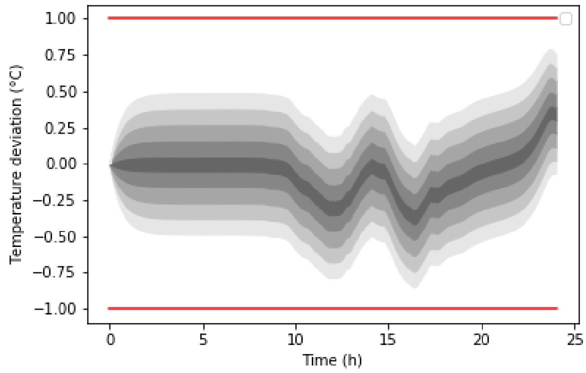
(a) Time evolution of power imbalance



(b) Time evolution of $\bar{u}^{(k)}$ for both types of devices



(c) Quantile plot of the temperature deviations of individual water heaters $\left(X^{(1,i,\infty)} - x^{\mathrm{ref},(1,i)}\right)_{1 \le i \le N_1}$



(d) Quantile plot of the temperature deviations of individual heat pumps $\left(X^{(2,i,\infty)} - x^{\mathrm{ref},(2,i)}\right)_{1 \le i \le N_2}$

Figure 3: Evolution of the system (1 scenario of solar irradiance, quantiles computed within each population class)

5.3.2 Numerical illustration of the convergence of the coordination signal to the limiting coordination signal

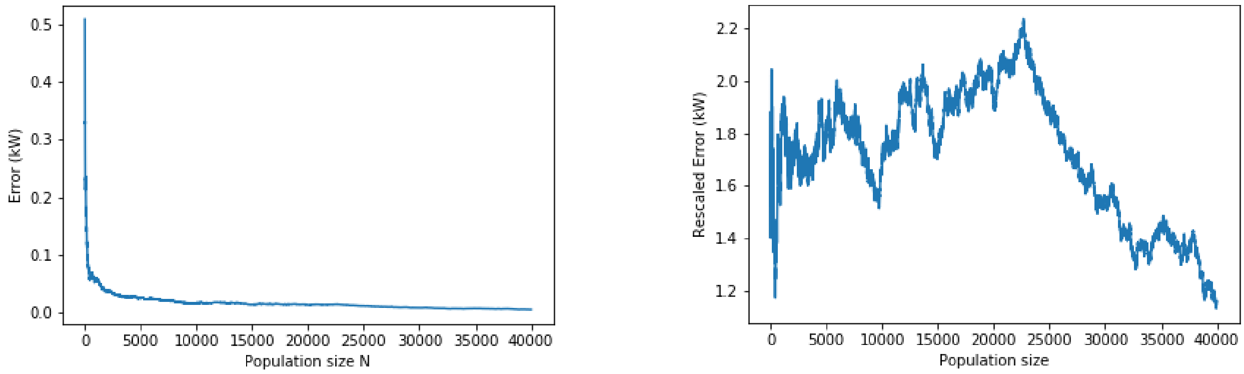
We plot the error between the real and limiting coordination signals $||\bar{v}^{(N)} - \bar{v}^{(\infty)}||_{\mathcal{H}}$ as a function of the population size N and conditionally to one scenario of $\mathrm{P}^{\mathrm{sun}}$ in Figure 4a and the rescaled error $\sqrt{N}||\bar{v}^{(N)} - \bar{v}^{(\infty)}||_{\mathcal{H}}$ as a function of the population size N (conditionally to one scenario of $\mathrm{P}^{\mathrm{sun}}$) 4b, to empirically illustrate the convergence of $\bar{v}^{(N)}$ to $\bar{v}^{(\infty)}$ at speed $O\left(\frac{1}{\sqrt{N}}\right)$ when the population size N goes to infinity. .

5.3.3 Impact of relative population sizes

We consider the relative sizes of the population given in Table 4 without modifying other parameters of the problem.

Table 4: Relative sizes of populations of water heaters and heat pumps

Scenario	Proportion of water heaters π_1	Proportion of heat pumps π_2
Case 1	50 %	50 %
Case 2	95 %	5 %
Case 3	5 %	95 %



(a) Error $\|\bar{v}^{(N)} - \bar{v}^{(\infty)}\|_{\mathcal{H}}$ as a function of population size N (conditionally to one scenario of $\mathrm{P}^{\mathrm{sun}}$)

(b) Rescaled error $\sqrt{N}\|\bar{v}^{(N)} - \bar{v}^{(\infty)}\|_{\mathcal{H}}$ as a function of population size N (conditionally to one scenario of $\mathrm{P}^{\mathrm{sun}}$)

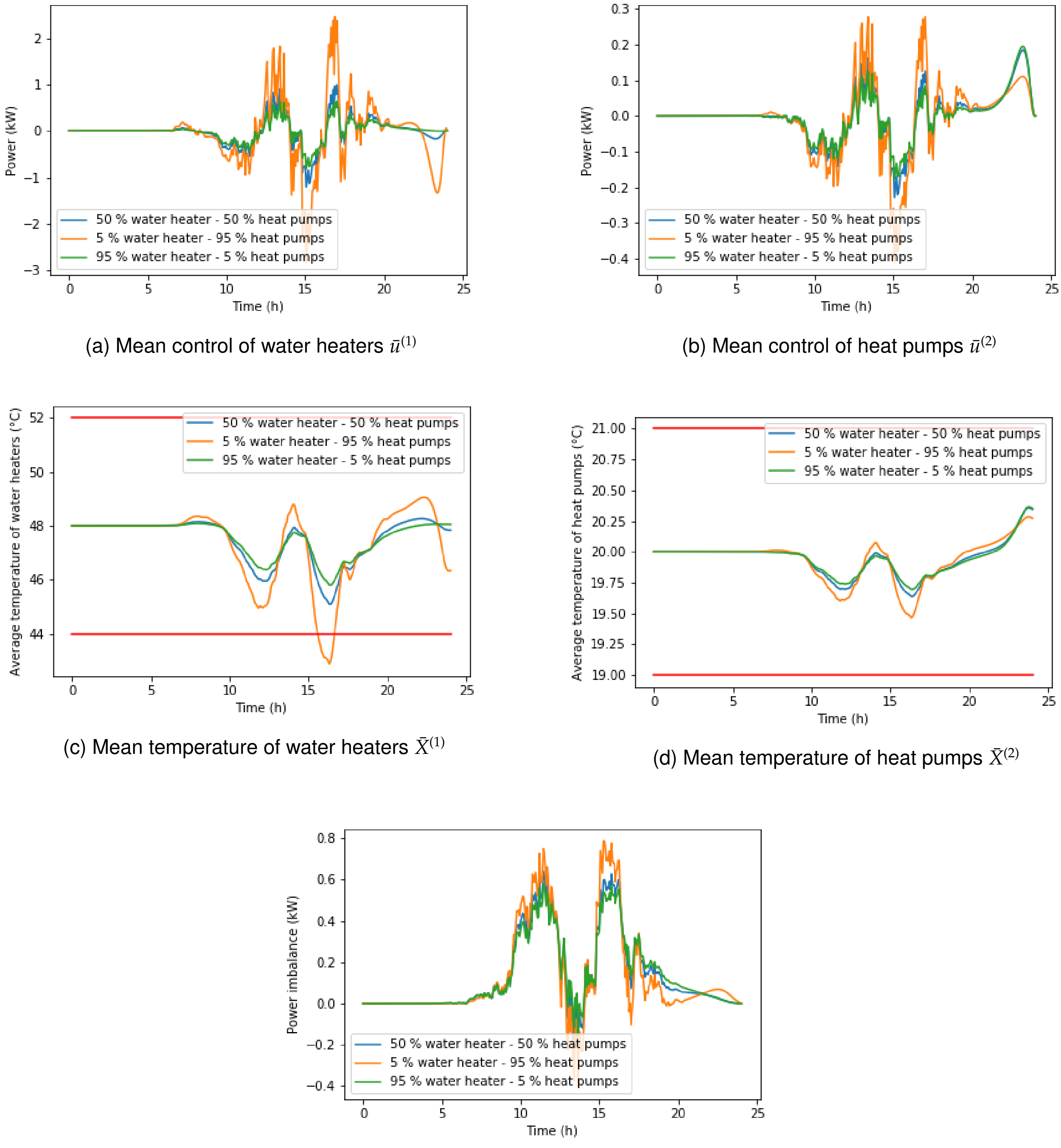
Figure 4: Convergence of the coordination signal to the limiting coordination signal in the limit of large populations

The evolution of the controlled power imbalance in the three cases is represented in Figure 5e. We observe a similar magnitude of this signal in the three cases considered. However, whenever the relative size of water heaters population is small (5%), the magnitudes of average controls of both types of population increase (see Figures 5a and 5b). Besides, the temperature of both types of TCLs vary more, and in the case of water heaters, the temperature may even go outside of the range defined by the deadband temperature (see Figures 5c and 5d). This may be explained intuitively. Water heaters have more capabilities to provide power without violating their operational constraints than heat pump do. As a result, water heater provide more power than heat pumps (see Figures 5a and 5b). Hence, when the relative population size of water heaters is reduced, the overall system has a smaller capability of providing and absorbing power. As a result, individual devices of both types are more solicited. These experiments show that there is a trade-off between ensuring individual constraints (power levels, temperature) and global power balance. Appropriate tuning of the parameters of the cost function is therefore required.

6 Online decentralized control scheme with minimal telecommunication

We go back to the setting of Section 4, with time dependent coefficients and with a non-quadratic loss function $\mathcal{L}$. We have developed a decomposition and a mean-field approximation which allow to solve approximately the stochastic control with limited telecommunications. In the control architecture developed, a coordinator solves the so-called limiting coordination problem (4.1) which allows him to compute the limiting coordination signal $\bar{v}^{(\infty)} \in \mathcal{H}$. This signal is sent to all agents and used as input parameter of the limiting individual problems 4.4. In particular, the (conditional) distribution of the limiting coordination signal $\bar{v}^{(\infty)}$ is required a priori by the agents to solve their individual problems. This raises two issues: sending the distribution of the limiting coordination signal $\bar{v}^{(\infty)}$ is costly and conceptually complex from a telecommunication point of view, and this information needs to be stored locally by agents, which requires heavy memory needs. To tackle these issues, we show that the affine-linear structure of the limiting individual problems allows to compute the current value of the control using a simpler coordination signal. This results in the same control as one would obtain by solving the limiting coordination and individual problems, without additional error. We describe an online control scheme where, at each instant, a coordinator sends the current best estimation (conditional expectation) of the limiting coordination signal on the remaining time horizon. Replacing the distribution of the limiting coordination signal by its current best estimation in the limiting individual problems allows to compute the current value of the control variables using a diagonal scheme. This allows a decentralized architecture with one-way real-time communication of a simple and agent-independent coordination signal from the coordinator to individual consumers, see Figure 6[4]. Let us give more details.

[4] Icons made by Freepik and Smashicons from www.flaticon.com



(a) Mean control of water heaters $\bar{u}^{(1)}$

(b) Mean control of heat pumps $\bar{u}^{(2)}$

(c) Mean temperature of water heaters $\bar{X}^{(1)}$

(d) Mean temperature of heat pumps $\bar{X}^{(2)}$

(e) Power imbalance

Figure 5: Impact of relative sizes of populations of TCLs for 1 scenario of $\mathtt{P}^{\mathrm{sun}}$ (means computed over the population)

Let us adopt the point of view of agent $i \in [N_k]$ of class $k \in [M]$. Consider its limiting individual problem:

$$\begin{cases} X_t = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s + \beta_s^{(k)} X_s + \gamma_s^{(k,i)}\right) \mathrm{d}s, \\ Y_t = \mathbb{E}_t \left[\rho^{(k)} \left(X_T - x_T^{\mathtt{f},(k)}\right) + \int_t^T \left(\beta_s^{(k)} Y_s + \nu_s^{(k)} \left(X_s - x_s^{\mathtt{ref},(k)}\right)\right) \mathrm{d}s\right], \\ \mu_t^{(k)} \left(u_t - u_t^{\mathtt{ref},(k)}\right) + \bar{v}_t^{(\infty)} + \alpha_t^{(k)} Y_t = 0, \end{cases} \tag{6.1}$$

where the limiting coordination signal $\bar{v}^{(\infty)} \in \mathcal{H}$ is defined in (4.3) and computed by the coordinator by solving the

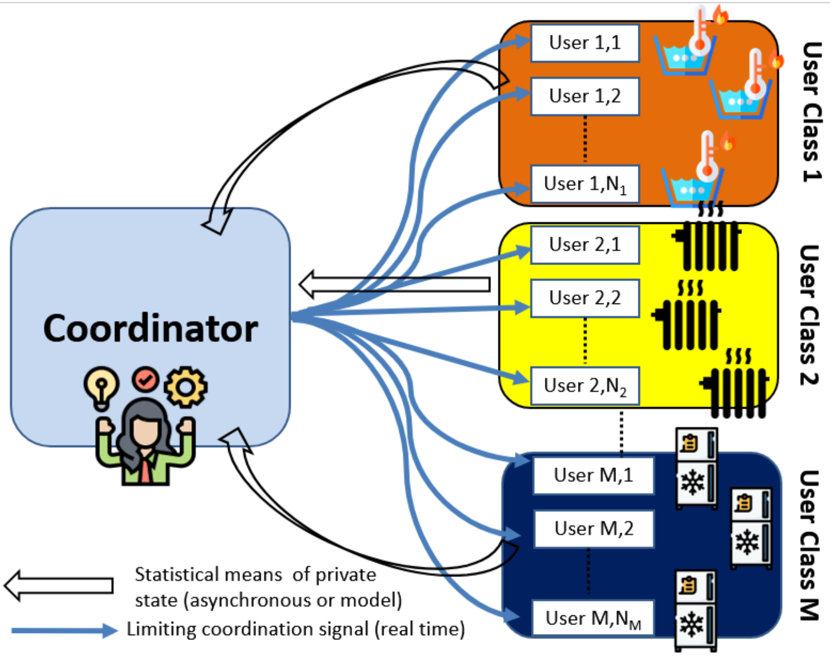


Figure 6: Coordination mechanism

limiting coordination problem (4.1). We recall that the solution of the individual problem of agent i of class k is denoted by $(u^{(k,i,\infty)}, X^{(k,i,\infty)}, Y^{(k,i,\infty)})$.

Introduce for $t \in [0, T]$ the online coordination signal at time t, denoted by $\bar{v}^{(\infty,t)}$ and defined by:

$$\bar{v}_\tau^{(\infty,t)} := \mathbb{E}_t\left[\bar{v}_\tau^{(\infty)}\right] = \mathbb{E}_t\left[\mathcal{L}'_x\left(\tau, \sum_{l=1}^{M} \pi^{(l)} \bar{u}_\tau^{(l,\infty)} + \bar{\mathrm{P}}_\tau^{\mathtt{load}} - \mathrm{P}_\tau^{\mathtt{prod}}\right)\right]. \tag{6.2}$$

As $\bar{v}^{(\infty)}$ is $\mathbb{G}$-progressively measurable and as $\mathbb{G}$ is immersed in $\mathbb{F}$, we also have:

$$\bar{v}_\tau^{(\infty,t)} = \mathbb{E}\left[\mathbb{E}\left[\bar{v}_\tau^{(\infty)} | \mathcal{G}_T\right] | \mathcal{F}_t\right] = \mathbb{E}\left[\bar{v}_\tau^{(\infty)} | \mathcal{G}_t\right].$$

In particular, $\bar{v}^{(\infty,t)}$ is a $\mathcal{G}_t$-measurable function of time which can be fully computed by the coordinator (which observes $\mathcal{G}_t$ only at time t). Let us consider the point of view of the agent i of class k at a fixed time $t \in [0, T]$, which aims at computing the current value of its control $u_t^{(k,i)}$. Informally, using the linearity of the conditional expectation and of (6.1), replacing $\bar{v}^{(\infty)}$ by its conditional expectation $\bar{v}^{(\infty,t)}$, we can heuristically justify that $(\mathbb{E}_t\left[u^{(k,i,\infty)}\right], \mathbb{E}_t\left[X^{(k,i,\infty)}\right], \mathbb{E}_t\left[Y^{(k,i,\infty)}\right])$ is solution of the following linear FBSDE with unknowns $(u, X, Y) \in \mathcal{H} \times \mathcal{S} \times \mathcal{S}$:

$$\begin{cases} X_\tau = x_0^{(k,i)} + \int_0^\tau \left(\alpha_s^{(k)} u_s + \beta_s^{(k)} X_s + \mathbb{E}_t\left[\gamma_s^{(k,i)}\right]\right) \mathrm{d}s \\ Y_\tau = \rho^{(k)}\left(X_T - \mathbb{E}_t\left[x_T^{\mathtt{f},(k)}\right]\right) + \int_\tau^T \left(\beta_s^{(k)} Y_s + \nu_s^{(k)}\left(X_s - \mathbb{E}_t\left[x_s^{\mathtt{ref},(k)}\right]\right)\right) \mathrm{d}s \\ \mu_\tau^{(k)}\left(u_\tau - \mathbb{E}_t\left[u_\tau^{\mathtt{ref},(k)}\right]\right) + \bar{v}_\tau^{(\infty,t)} + \alpha_\tau^{(k)} Y_\tau = 0. \end{cases} \tag{6.3}$$

Using the "diagonal" identity

$$(u_t^{(k,i,\infty)}, X_t^{(k,i,\infty)}, Y_t^{(k,i,\infty)}) = (\mathbb{E}_t\left[u_t^{(k,i,\infty)}\right], \mathbb{E}_t\left[X_t^{(k,i,\infty)}\right], \mathbb{E}_t\left[Y_t^{(k,i,\infty)}\right]), \tag{6.4}$$

we can recover the current value of the control $u^{(k,i)}$ of the agent at time t. This procedure can be repeated for each time t: the coordinator solves the coordination problem, evaluates the online coordination signal $\bar{v}^{(\infty,t)}$ on the remaining time horizon and sends it to all agents. The agents can then compute the current value of their control by solving a one-dimensional affine-linear FBSDE without additional approximation error. Such a problem is easy to solve, see Theorem 5.2.

The above discussion justifies the following decentralized control scheme 1.

Algorithm 1 Decentralized control scheme

Inputs: Time grid $(\tau_0, ..., \tau_{N_T})$.
for $j = 0, ..., N_T$ **do**
  Wait for $t = \tau_j$.
  Aggregator observes the common noise $\mathcal{G}_{\tau_j}$, computes the online coordination signal at time τ_j $(\bar{v}_\tau^{(\infty,\tau_j)})_{\tau_j \le \tau \le T}$ given in (6.2) and sends it to all agents. This coordination signal is a $\mathcal{G}_{\tau_j}$ measurable function of time.
  Each agent solves its limiting online individual problem (6.3) to get its optimal control $u_{\tau_j}^{(k,i,\infty)} = u_{\tau_j}^{(k,i,\infty,\tau_j)}$ by the diagonal identity (6.4).
  Each agent implements control $u_{\tau_j}^{(k,i,\infty)}$ for its storage system at time τ_j.
end for

6.1 Discussion on the decentralized control scheme

We can make the following remarks on this scheme.

1. **Fast computation of the coordination signal is possible with reasonable computational resource.** The limiting coordination problem (4.1) is a M-dimensional FBSDE, equivalent to a M-dimensional control problem. The control problem is relatively easy to solve in a linear-quadratic setting, using similar arguments as in Section 5.2, and it can be solved using numerical methods in other cases. Computation of the parameters of this problem is easy as well, under some assumption like affine-linear stochastic processes.

2. **Fast computation of the individual controls are possible by agents equipped with limited computational resources.** The online limiting individual problems at time t (6.3) are linear one-dimensional FBSDEs, hence particularly easy to solve, see Section 5.2. Considering deterministic coefficients or stochastic processes with affine-linear drift can make computations even easier.

3. **The parameters of the problems solved by the agents and the coordinator are locally available.** Indeed, the parameters of the limiting coordination problem at time t can be computed by an aggregator only observing the common information $\mathbb{G}$. The parameters of the online limiting individual problems of agent (k, i) at time t are all available locally for agent (k, i). This includes the shared information $\mathcal{G}_t$, the online limiting coordination signal received, and individual parameters of the energy storage system of agent (k, i). In particular, the computation of conditional expectations of the parameters of individual energy storage systems is simplified by our assumption of conditional independence of these parameters.

4. **Limited telecommunication is required.** A single online coordination signal at time t $(\bar{v}_\tau^{(\infty,t)})\tau \in [0, T]$ is sent to agents by the coordinator, so that no specific routing is needed. This signal is a one dimensional $\mathcal{G}_t$-measurable function of time, hence it encoding is easy, for instance by discretization/interpolation or by regression against a function basis (like Fourier). No real-time communication from agents to the coordinator is required.

6.2 On the privacy of individual users habits

In order for an aggregator to come up with good stochastic models of the empirical averages of the class parameters $(\bar{\mathrm{P}}^{\mathtt{load},(k,N)}, \bar{\gamma}^{(k,N)}, \bar{u}^{\mathtt{ref},(k,N)}, \bar{x}^{\mathtt{ref},(k,N)}, \bar{x}_T^{\mathtt{f},(k,N)})_{k \in [M]}$, one may imagine that the aggregator is given some historical realization of these processes. As these processes are aggregates of individual data of consumers, which may be subject to some privacy requirements, one can use the Secure Multiparty Computation (SMC) technique in [Yao86] in order to deal with privacy concerns. Indeed, this technique would allow the coordinator to compute the empirical averages of the class parameters, while guaranteeing that the values of the parameters of individual agents remain unknown to him. This method has already been used in the context of energy management in [Jac+19].

7 Proofs

7.1 Proof of Theorem 2.1

The existence and uniqueness of an optimal control are proved using standard arguments of functional analysis. We give main arguments and leave full details to the reader. The convexity directly stems from the linearity of the dynamic and the quadratic/convex functions in the definition of $\mathcal{J}$. In addition, the strong convexity of $\mathcal{J}$ comes from the uniform lower bound on μ. It directly yields the coercivity of $\mathcal{J}$. As $\mathcal{J}$ is additionally continuous, we get the existence of a minimizer of $\mathcal{J}$ from [Bre10, Corollary 3.23, pp.71]. This minimizer is unique from the strict convexity of $\mathcal{J}$. We denote it by $(u^{(k,i,N)})_{k\in[M],i\in[N_k]} \in \mathcal{H}$.

The characterization of optimality is proved applying the stochastic Pontryagin principle. However, our setting of optimal control of ODE with non-Markovian coefficients in general filtrations differs from standard references: see [Yon99] for the case of Brownian filtrations, see [CD18, pp. 543-552, Volume I] when incorporating McKean-Vlasov terms. The closest reference to our setting is presumably [Cad02] in the (Markovian) SDE case with jumps, under different integrability conditions. This motivates us to give a proof of our result in our specific setting.

Let $u = (u^{(k,i)})_{k\in[M],i\in[N_k]} \in \mathcal{H}$. By our integrability assumptions, there is existence and uniqueness of $X = (X^{(k,i)})_{k\in[M],i\in[N_k]} \in \mathcal{S}$ solution of the ODE:

$$\forall k \in [M], \forall i \in [N_k], \quad X_t^{(k,i)} = x_0^{(k,i)} + \int_0^t \left(\alpha_s^{(k)} u_s^{(k,i)} + \beta_s^{(k)} X_s^{(k,i)} + \gamma_s^{(k,i)}\right) \mathrm{d}s. \tag{7.1}$$

Now, let us consider the Backward Stochastic Differential Equation in $(Y, \tilde{M}) := (Y^{(k,i)}, \tilde{M}^{(k,i)})_{k\in[M],i\in[N_k]} \in \mathcal{S}\times\mathcal{S}$ with $\tilde{M}$ a square integrable martingale vanishing at $t = 0$:

$$\forall k \in [M], \forall i \in [N_k], \begin{cases} -\mathrm{d}Y_t^{(k,i)} = \left(\beta_t^{(k)} Y_t^{(k,i)} + \nu_t^{(k)}(X_t^{(k,i)} - x_t^{\mathtt{ref},(k,i)})\right)\mathrm{d}t - \mathrm{d}\tilde{M}_t^{(k,i)}, \\ Y_T^{(k,i)} = \rho^{(k)}\left(X_T^{(k,i)} - x_T^{\mathtt{f},(k,i)}\right). \end{cases}$$

It is an affine-linear BSDE in a general filtration, and our boundedness and integrability assumptions on its coefficients ensure existence and uniqueness of its solution, see [EPQ97, Theorem 5.1, p. 54].

Now, the arguments for proving the Gateaux-differentiability of $\mathcal{J}$ are standard and follow the ones in [CD18, pp. 543-548, Volume I]: we show Gateaux-differentiability of the state variable and of the cost functional successively. Define the application $\phi_X : u := (u^{(k,i)})_{k\in[M],i\in[N_k]} \in \mathcal{H} \mapsto X^u := (X^{(k,i)})_{k\in[M],i\in[N_k]} \in \mathcal{S}$ by (7.1). Then ϕ_X is Gateaux differentiable and its Gateaux derivative at $u := (u^{(k,i)})_{k\in[M],i\in[N_k]}$ in direction $v := (v^{(k,i)})_{k\in[M],i\in[N_k]}$ is given by $\left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}X^{u+\varepsilon v}\right)_{|\varepsilon=0} = \dot{X}^v := (\dot{X}^{v,(k,i)})_{k\in[M],i\in[N_k]}$ with:

$$\dot{X}_t^{v,(k,i)} = \int_0^t \left(\alpha_s^{(k)} v_s^{(k,i)} + \beta_s^{(k)} \dot{X}_s^{v,(k,i)}\right) \mathrm{d}s.$$

This can be proved following arguments of the proof of [CD18, Lemma 6.10, pp.544-545, Volume I]. Now, following arguments in [CD18, pp. 546-548, Volume I], we get that $\mathcal{J}$ is Gateaux-differentiable and its Gateaux derivative at u in direction v is given by $\left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathcal{J}(u+\varepsilon v)\right)_{|\varepsilon=0} = \dot{\mathcal{J}}(u, v)$ where:

$$\begin{aligned}\dot{\mathcal{J}}(u, v) =& \mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k}\int_0^T \left(\mu_t^{(k)}\left(u_t^{(k,i)} - u_t^{\mathtt{ref},(k,i)}\right)v_t^{(k,i)} + \nu_t^{(k)}\left(X_t^{(k,i)} - x_t^{\mathtt{ref},(k,i)}\right)\dot{X}_t^{v,(k,i)}\right)\mathrm{d}t\right] \\ &+ \mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k}\left(\rho^{(k)}\left(X_T^{(k,i)} - x_T^{\mathtt{f},(k,i)}\right)\dot{X}_T^{v,(k,i)} + \int_0^T \mathcal{L}'_x\left(t, \frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\left(u_t^{(l,j)} + \mathrm{P}_t^{\mathtt{load},(l,j)}\right) - \mathrm{P}_t^{\mathtt{prod}}\right)v_t^{(k,i)}\mathrm{d}t\right)\right].\end{aligned}$$

Then, applying Integration by Parts Formula in [Pro03, Corollary 2, p. 68] to $Y \cdot \dot{X}^v := \sum_{k=1}^{M}\sum_{i=1}^{N_k} Y^{(k,i)}\dot{X}^{v,(k,i)}$ between $t = 0$ and $t = T$ yields, using $\dot{X}_0^v = 0$ and $Y_T^{(k,i)} = \rho^{(k)}\left(X_T^{(k,i)} - x_T^{\mathtt{f},(k,i)}\right)$, we finally obtain the following expression for the Gateaux derivative of $\mathcal{J}$ at u in direction v:

$$\dot{\mathcal{J}}(u, v) = \mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k}\int_0^T \left\{\mu_t^{(k)}\left(u_t^{(k,i)} - u_t^{\mathtt{ref},(k,i)}\right) + \mathcal{L}'_x\left(t, \frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\left(u_t^{(l,j)} + \mathrm{P}_t^{\mathtt{load},(l,j)}\right) - \mathrm{P}_t^{\mathtt{prod}}\right) + \alpha_t^{(k)}Y_t^{(k,i)}\right\}v_t^{(k,i)}\mathrm{d}t\right]. \tag{7.2}$$

By convexity and differentiability of $\mathcal{J}$ and by uniqueness of its minimizer, $(u^{(k,i,N)})_{k\in[M],i\in[N_k]} \in \mathcal{H}$ is also the unique critical point of $\mathcal{J}$. Combining this with the expression of the Gateaux derivative of $\mathcal{J}$, we get that the term inside the brackets in (7.2) is 0 for all t. Therefore, $(u^{(k,i,N)}, X^{(k,i,N)}, Y^{(k,i,N)})_{k\in[M],i\in[N_k]} \in \mathcal{H}\times\mathcal{S}\times\mathcal{S}$ is the unique solution of the FBSDE (2.4). □

7.2 Proof of Proposition 3.1

Using Theorem 2.1 and the definition of the empirical mean processes, one can directly show that the empirical mean processes are solution of (3.1). This FBSDE fully characterizes the solution in $(u^{(k)})_{k\in[M]} \in \mathcal{H}^M$ of the following stochastic control problem:

$$
\begin{aligned}
\min_{(u^{(k)})_{k\in[M]}\in\mathcal{H}^M} \quad & \mathbb{E}\left[\sum_{k=1}^{M}\pi^{(k)}\left\{\int_0^T\left(\frac{\mu_t^{(k)}}{2}\left(u_t^{(k)}-\bar{u}_t^{\mathtt{ref},(k,N)}\right)^2+\frac{\nu_t^{(k)}}{2}\left(X_t^{(k)}-\bar{x}_t^{\mathtt{ref},(k,N)}\right)^2\right)\mathrm{d}t+\frac{\rho^{(k)}}{2}\left(X_T^{(k)}-\bar{x}_T^{\mathtt{f},(k,N)}\right)^2\right\}\right] \\
& +\mathbb{E}\left[\int_0^T\mathcal{L}_t\left(\sum_{l=1}^{M}\pi^{(l)}u_t^{(l)}+\bar{\mathrm{P}}_t^{\mathtt{load},(N)}-\mathrm{P}_t^{\mathtt{prod}}\right)\mathrm{d}t\right], \qquad (7.3)\\
s.t. \quad & X_t^{(k)}=\frac{1}{N_k}\sum_{j=1}^{N_k}x_0^{(k,j)}+\int_0^t\left(\alpha_s^{(k)}u_s^{(k)}+\beta_s^{(k)}X_s^{(k)}+\bar{\gamma}_s^{(k,N)}\right)\mathrm{d}s, \quad \forall k\in[M].
\end{aligned}
$$

This results from applying Pontryagin's principle, up to scaling of the k-th adjoint variable by $\frac{1}{\pi^{(k)}}$, and by similar arguments as the ones used in the proof of Theorem 2.1. The uniqueness of the solution of the control problem, and hence of the solution of the FBSDE (3.1) from similar arguments as the ones used in the proof of Theorem 2.1. □

7.3 Proof of Proposition 4.3

By independence of $(\mathrm{P}_t^{\mathtt{load},(l,j)})_{1\le l\le M,1\le j\le N_l}$ conditionally to $\mathcal{G}_T$:

$$
\begin{aligned}
\mathbb{V}\mathrm{ar}\left[\frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\mathrm{P}_t^{\mathtt{load},(l,j)}|\mathcal{G}_T\right] &= \frac{1}{N^2}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\mathbb{V}\mathrm{ar}\left[\mathrm{P}_t^{\mathtt{load},(l,j)}|\mathcal{G}_T\right],\\
\mathbb{E}\left[\frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\mathrm{P}_t^{\mathtt{load},(l,j)}|\mathcal{G}_T\right] &= \bar{\mathrm{P}}_t^{\mathtt{load}}.
\end{aligned}
$$

This yields:

$$
\mathbb{E}\left[(\bar{\mathrm{P}}_t^{\mathtt{load},(N)}-\bar{\mathrm{P}}_t^{\mathtt{load}})^2\right]=\mathbb{E}\left[\mathbb{V}\mathrm{ar}\left[\bar{\mathrm{P}}_t^{\mathtt{load},(N)}|\mathcal{G}_T\right]\right]=\frac{1}{N^2}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\mathbb{E}\left[\mathbb{V}\mathrm{ar}\left[\mathrm{P}_t^{\mathtt{load},(l,j)}|\mathcal{G}_T\right]\right]\le\frac{1}{N^2}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\mathbb{E}\left[\left(\mathrm{P}_t^{\mathtt{load},(l,j)}\right)^2\right].
$$

This yields, integrating over time and using the fact that all $\mathrm{P}^{\mathtt{load},(l,j)}$ are bounded in $\mathcal{H}$ by a constant independent from N:

$$
\|\bar{\mathrm{P}}^{\mathtt{load},(N)}-\bar{\mathrm{P}}^{\mathtt{load}}\|_{\mathcal{H}}^2\le\frac{C}{N}
$$

Similarly, we obtain the convergence in $\mathcal{H}$ of $\bar{\gamma}^{(k,N)}$ (resp. $\bar{u}^{\mathtt{ref},(k,N)}$, resp. $\bar{x}^{\mathtt{ref},(k,N)}$) to $\bar{\gamma}^{(k)}$ (resp. $\bar{u}^{\mathtt{ref},(k)}$, resp. $\bar{x}^{\mathtt{ref},(k)}$) at speed $\frac{1}{\sqrt{N_k}}$, and the convergence in $\mathbb{L}^2$ of $\bar{x}_T^{\mathtt{f},(k)}$ to $\bar{x}_T^{\mathtt{f},(k)}$ at speed $\frac{1}{\sqrt{N_k}}$. □

7.4 Proof of Theorem 4.5

Consider the following FBSDE with $\mathbb{G}$-progressively measurable coefficients:

$$\forall k\in[M],$$

$$
\begin{cases}
X_t^{(k)} = \bar{x}_0^{(k)} + \int_0^t \left(\alpha_s^{(k)} u_s^{(k)} + \beta_s^{(k)} X_s^{(k)} + \bar{\gamma}_s^{(k)}\right) \mathrm{d}s, \\
Y_t^{(k)} = \mathbb{E}\left[\rho^{(k)}\left(X_T^{(k)} - \bar{x}_T^{\mathtt{f},(k)}\right) + \int_t^T \left(\beta_s^{(k)} Y_s^{(k)} + \nu_s^{(k)}\left(X_s^{(k)} - \bar{x}_s^{\mathtt{ref},(k)}\right)\right) \mathrm{d}s | \mathcal{G}_t\right], \\
\mu_t^{(k)}\left(u_t^{(k)} - \bar{u}_t^{\mathtt{ref},(k)}\right) + \mathcal{L}_x'\left(t, \sum_{l=1}^M \pi^{(l)} u_t^{(l)} + \bar{\mathrm{P}}_t^{\mathtt{load}} - \mathrm{P}_t^{\mathtt{prod}}\right) + \alpha_t^{(k)} Y_t^{(k)} = 0.
\end{cases}
\tag{7.4}
$$

The above FBSDE is the optimality system associated to the following stochastic control problem considered in in $(\Omega, \mathcal{G}_T, \mathbb{G}, \mathbb{P})$:

$$
\begin{aligned}
\min_{(u^{(k)})_{1\le k\le M}\in\mathcal{H}} \quad & \mathbb{E}\left[\sum_{k=1}^M \pi^{(k)} \left\{\int_0^T \left(\frac{\mu_t^{(k)}}{2}\left(u_t^{(k)} - \bar{u}_t^{\mathtt{ref},(k)}\right)^2 + \frac{\nu_t^{(k)}}{2}\left(X_t^{(k)} - \bar{x}_t^{\mathtt{ref},(k)}\right)^2\right) \mathrm{d}t + \frac{\rho^{(k)}}{2}\left(X_T^{(k)} - \bar{x}_T^{\mathtt{f},(k)}\right)^2\right\}\right] \\
& + \mathbb{E}\left[\int_0^T \mathcal{L}_t\left(\sum_{l=1}^M \pi^{(l)} u_t^{(l)} + \bar{\mathrm{P}}_t^{\mathtt{load}} - \mathrm{P}_t^{\mathtt{prod}}\right) \mathrm{d}t\right], \\
s.t. \quad & X_t^{(k)} = \bar{x}_0^{(k)} + \int_0^t \left(\alpha_s^{(k)} u_s^{(k)} + \beta_s^{(k)} X_s^{(k)} + \bar{\gamma}_s^{(k)}\right) \mathrm{d}s, \quad \forall k \in [M].
\end{aligned}
$$

Our assumptions and [Bre10, Corollary 3.23, pp.71] show that the above problem has a unique solution $\tilde{u} \in \mathcal{H}_\mathbb{G}$ and therefore, the FBSDE (7.4) has a unique solution $(\tilde{u}, \tilde{X}, \tilde{Y}) \in \mathcal{H}_\mathbb{G} \times \mathcal{S}_\mathbb{G} \times \mathcal{S}_\mathbb{G}$. Now consider the $\mathbb{G}$-martingales:

$$
\tilde{M}_t^{(k)} := \mathbb{E}\left[\rho^{(k)}\left(\tilde{X}_T^{(k)} - \bar{x}_T^{\mathtt{f},(k)}\right) + \int_0^T \left(\beta_s^{(k)} \tilde{Y}_s^{(k)} + \nu_s^{(k)}\left(\tilde{X}_s^{(k)} - \bar{x}_s^{\mathtt{ref},(k)}\right)\right) \mathrm{d}s | \mathcal{G}_t\right].
$$

Noting that $\mathbb{G}$ is immersed in $\mathbb{F}$ (see [CD18, Definition 1.2, p. 5, Volume II]), for all $k \in [M]$, $\tilde{M}^{(k)}$ is a $\mathbb{G}$-square integrable martingale and therefore, by definition, it is a $\mathbb{F}$-square integrable martingale, so that:

$$
\tilde{M}_t^{(k)} := \mathbb{E}_t\left[\rho^{(k)}\left(\tilde{X}_T^{(k)} - \bar{x}_T^{\mathtt{f},(k)}\right) + \int_0^T \left(\beta_s^{(k)} \tilde{Y}_s^{(k)} + \nu_s^{(k)}\left(\tilde{X}_s^{(k)} - \bar{x}_s^{\mathtt{ref},(k)}\right)\right) \mathrm{d}s\right].
$$

Therefore, we have for all $k \in [M]$, by the previous two expressions of $\tilde{M}^{(k)}$ and using the fact that $\tilde{X}^{(k)}$, $\tilde{Y}^{(k)}$ and $\bar{x}^{\mathtt{ref},(k)}$ are $\mathbb{G}$ and $\mathbb{F}$ progressively measurable (as $\mathbb{G}$ is assumed immersed in $\mathbb{F}$):

$$
\begin{aligned}
\tilde{Y}_t^{(k)} &= \mathbb{E}\left[\rho^{(k)}\left(\tilde{X}_T^{(k)} - \bar{x}_T^{\mathtt{f},(k)}\right) + \int_t^T \left(\beta_s^{(k)} \tilde{Y}_s^{(k)} + \nu_s^{(k)}\left(\tilde{X}_s^{(k)} - \bar{x}_s^{\mathtt{ref},(k)}\right)\right) \mathrm{d}s | \mathcal{G}_t\right] \\
&= \tilde{M}_t^{(k)} - \int_0^t \left(\beta_s^{(k)} \tilde{Y}_s^{(k)} + \nu_s^{(k)}\left(\tilde{X}_s^{(k)} - \bar{x}_s^{\mathtt{ref},(k)}\right)\right) \mathrm{d}s \\
&= \mathbb{E}_t\left[\rho^{(k)}\left(\tilde{X}_T^{(k)} - \bar{x}_T^{\mathtt{f},(k)}\right) + \int_t^T \left(\beta_s^{(k)} \tilde{Y}_s^{(k)} + \nu_s^{(k)}\left(\tilde{X}_s^{(k)} - \bar{x}_s^{\mathtt{ref},(k)}\right)\right) \mathrm{d}s\right].
\end{aligned}
$$

Besides, $\tilde{u}$ and $\tilde{X}$ are $\mathbb{F}$-progressively measurable. Therefore, $(\tilde{u}, \tilde{X}, \tilde{Y})$ is also solution of the optimality system of the control problem considered with the filtration $\mathbb{F}$. By uniqueness of such a solution, we deduce that $(\bar{u}, \bar{X}, \bar{Y})$ coincides with $(\tilde{u}, \tilde{X}, \tilde{Y})$ and therefore $(\bar{u}, \bar{X}, \bar{Y})$ is $\mathbb{G}$-progressively measurable. □

7.5 Proof of Proposition 4.6

The uniqueness of solution of the above FBSDE arises from similar arguments as in the proof of Theorem 2.1.

Consider the function $\chi : [0,T] \times \mathbb{R}^M \times \mathbb{R} \times \mathbb{R}^M \times \mathbb{R}^M \mapsto \mathbb{R}$:

$$
\chi : (t, u, v, y, u^{\mathtt{ref}}) \mapsto \sum_{k=1}^M \left(\frac{\pi^{(k)}\mu^{(k)}}{2}(u_k - u^{\mathtt{ref},(k)})^2 + \pi^{(k)}\alpha_t^{(k)} y_k u_k\right) + \mathcal{L}_t\left(\sum_{l=1}^M \pi^{(l)} u_i + v\right).
$$

It is straightforward to observe that χ is twice continuously differentiable in $(u, v, y, u^{\mathtt{ref}})$.

For any $(t, v, y, u^{\mathtt{ref}}) \in [0,T] \times \mathbb{R} \times \mathbb{R}^M \times \mathbb{R}^M$, $u \mapsto \chi(t, u, v, y, u^{\mathtt{ref}})$ is $\min_k\{\pi_k \mu^{(k)}\}$-strongly convex, and as such, this function admits a unique minimizer and its Hessian is positive semi-definite. Besides, its Hessian is invertible with inverse bounded by $\frac{1}{\min_k\{\pi_k \mu^{(k)}\}}$.

By the implicit function theorem, this directly implies, for any $(t, v, y, u^{\mathrm{ref}}) \in [0,T] \times \mathbb{R} \times \mathbb{R}^M \times \mathbb{R}^M$, the equation in $u \in \mathbb{R}^M$

$$\nabla_u \chi(t, u, v, y, u^{\mathrm{ref}}) = 0$$

has a unique solution $u = \tilde{u}(t, v, y, u^{\mathrm{ref}})$ with $\tilde{u}$ continuously differentiable in (v, y, u^{ref}). Besides, we have:

$$\begin{aligned}
\nabla_v \tilde{u}(t, v, y, u^{\mathrm{ref}}) &= -\left(\nabla^2_{uu}\chi(t, \tilde{u}(t, v, y, u^{\mathrm{ref}}), v, y, u^{\mathrm{ref}})\right)^{-1}\left(\nabla^2_{u,v}\chi(t, \tilde{u}(t, v, y, u^{\mathrm{ref}}), v, y, u^{\mathrm{ref}})\right), \\
\nabla_y \tilde{u}(t, v, y, u^{\mathrm{ref}}) &= -\left(\nabla^2_{uu}\chi(t, \tilde{u}(t, v, y, u^{\mathrm{ref}}), v, y, u^{\mathrm{ref}})\right)^{-1}\left(\nabla^2_{u,y}\chi(t, \tilde{u}(t, v, y, u^{\mathrm{ref}}), v, y, u^{\mathrm{ref}})\right), \\
\nabla_{u^{\mathrm{ref}}} \tilde{u}(t, v, y, u^{\mathrm{ref}}) &= -\left(\nabla^2_{uu}\chi(t, \tilde{u}(t, v, y, u^{\mathrm{ref}}), v, y, u^{\mathrm{ref}})\right)^{-1}\left(\nabla^2_{u,u^{\mathrm{ref}}}\chi(t, \tilde{u}(t, v, y, u^{\mathrm{ref}}), v, y, u^{\mathrm{ref}})\right).
\end{aligned}$$

Using the bound $\|\left(\nabla^2_{uu}\chi(t, \tilde{u}(t, v, y, u^{\mathrm{ref}}), v, y, u^{\mathrm{ref}})\right)^{-1}\| \leq \frac{1}{\min_k\{\pi_k \mu^{(k)}\}}$ and as the second order derivative of $\mathcal{L}$ with respect to x is uniformly bounded, we obtain:

$$\|\nabla_v \tilde{u}(t, v, y, u^{\mathrm{ref}})\| + \|\nabla_y \tilde{u}(t, v, y, u^{\mathrm{ref}})\| + \|\nabla_{u^{\mathrm{ref}}} \tilde{u}(t, v, y, u^{\mathrm{ref}})\| \leq \frac{C}{\min_k\{\pi_k \mu_t^{(k)}\}}.$$

Then there exists a constant C which grows like $\frac{1}{\min_{1 \leq k \leq M, t \in [0,T]}\{\pi_k \mu_t^{(k)}\}}$ such that, for any $y^1, y^2 \in \mathbb{R}^M$, $v^1, v^2 \in \mathbb{R}$ and $u^{\mathrm{ref},1}, u^{\mathrm{ref},2} \in \mathbb{R}^M$, we have:

$$\|\tilde{u}(t, v^1, y^1, u^{\mathrm{ref},1}) - \tilde{u}(t, v^2, y^2, u^{\mathrm{ref},2})\|_{\mathbb{R}^M} \leq C\left(|v^1 - v^2| + \|y^1 - y^2\|_{\mathbb{R}^M} + \|u^{\mathrm{ref},1} - u^{\mathrm{ref},2}\|_{\mathbb{R}^M}\right).$$

This implies, for $\theta^1 = (x^1, v^1, w^1, u^{\mathrm{ref},1}, x^{\mathrm{ref},1}, x_T^{\mathrm{f},1})$ and $\theta^2 = (x^2, v^2, w^2, u^{\mathrm{ref},2}, x^{\mathrm{ref},2}, x_T^{\mathrm{f},2})$ in $\mathbb{R}^M \times \mathcal{H} \times \mathcal{H}(\mathbb{R}^M) \times \mathcal{H}(\mathbb{R}^M) \times \mathbb{L}^2_T(\mathbb{R}^M)$:

$$\|u^{\theta^1} - u^{\theta^2}\|_{\mathcal{H}} \leq C\left(\|v^1 - v^2\|_{\mathcal{H}} + \|Y^{\theta^1} - Y^{\theta^2}\|_{\mathcal{H}} + \|u^{\mathrm{ref},1} - u^{\mathrm{ref},2}\|_{\mathcal{H}}\right).$$

In the following, C_T denotes a constant depending on the input parameters of the problem and depending continuously on T, with finite limit when $T \to 0$.

Applying Gronwall's lemma to the state equations, we obtain:

$$\|X^{\theta^1} - X^{\theta^2}\|_{\mathcal{S}} \leq C_T\left(\|x^1 - x^2\|_{\mathbb{R}^M} + \|u^{\theta^1} - u^{\theta^2}\|_{\mathcal{H}} + \|w^1 - w^2\|_{\mathcal{H}}\right).$$

Applying Gronwall's lemma and Cauchy-Schwartz inequality to the adjoint equations yields:

$$\|Y^{\theta^1} - Y^{\theta^2}\|_{\mathcal{S}} \leq C_T\left(\|X^{\theta^1} - X^{\theta^2}\|_{\mathcal{S}} + \|x^{\mathrm{ref},1} - x^{\mathrm{ref},2}\|_{\mathcal{H}} + \|x_T^{\mathrm{f},1} - x_T^{\mathrm{f},2}\|_{\mathbb{L}^2}\right).$$

Combining the previous inequalities, we get:

$$\begin{aligned}
\|Y^{\theta^1} - Y^{\theta^2}\|_{\mathcal{S}} &\leq C_T\left(\|Y^{\theta^1} - Y^{\theta^2}\|_{\mathcal{H}} + \|\theta^1 - \theta^2\|\right) \\
&\leq C_T\left(\sqrt{T}\|Y^{\theta^1} - Y^{\theta^2}\|_{\mathcal{S}} + \|\theta^1 - \theta^2\|\right).
\end{aligned}$$

Then, using the fact that C_T is bounded for T small, we obtain, for any T small enough, so that $C_T\sqrt{T} < 1$:

$$\|Y^{\theta^1} - Y^{\theta^2}\|_{\mathcal{S}} \leq C_T\|\theta^1 - \theta^2\|.$$

Combining with the above estimations, this finally yields:

$$\|(\bar{u}^{\theta^1} - \bar{u}^{\theta^2}, \bar{X}^{\theta^1} - \bar{X}^{\theta^2}, \bar{Y}^{\theta^1} - \bar{Y}^{\theta^2})\|_{\mathcal{H}} \leq C_T\|\theta^1 - \theta^2\|.$$

□

7.6 Proof of Corollary 4.8

Apply Proposition 4.6 to $x^1 = \left(\bar{x}_0^{(k)}\right)_{1\le k\le M}$, $x^2 = \left(\frac{1}{N_k}\sum_{j=1}^{N_k} x_0^{(k,j)}\right)_{1\le k\le M}$, $v^1 = \bar{\mathrm{P}}^{\mathrm{load}} - \mathrm{P}^{\mathrm{prod}}$, $v^2 = \bar{\mathrm{P}}^{\mathrm{load},(N)} - \mathrm{P}^{\mathrm{prod}}$, $w^1 = \left(\bar{y}^{(k)}\right)_{1\le k\le M}$, $w^2 = \left(\bar{y}^{(k,N)}\right)_{1\le k\le M}$, $u^{\mathrm{ref},1} = (\bar{u}^{\mathrm{ref},(k)})_{1\le k\le M}$, $u^{\mathrm{ref},2} = (\bar{u}^{\mathrm{ref},(k,N)})_{1\le k\le M}$, $x^{\mathrm{ref},1} = (\bar{x}^{\mathrm{ref},(k)})_{1\le k\le M}$, , $x^{\mathrm{ref},2} = (\bar{x}^{\mathrm{ref},(k,N)})_{1\le k\le M}$, , $x_T^{\mathrm{f},1} = (\bar{x}_T^{\mathrm{f},(k)})_{1\le k\le M}$, $x_T^{\mathrm{f},1} = (\bar{x}_T^{\mathrm{f},(k)})_{1\le k\le M}$. Then use Proposition 4.3 and the assumption $\pi^{(k)} \ge \eta > 0$ for all $1 \le k \le N$ to conclude. □

7.7 Proof of Theorem 4.12

By applying the previous proposition to $v = \bar{v}^{(N)} = \frac{1}{N}\sum_{l=1}^{M}\sum_{j=1}^{N_l}\left(u^{(l,j,N)} + \mathrm{P}^{\mathrm{load},(l,j)}\right) - \mathrm{P}^{\mathrm{prod}}$ and $v' = \bar{v}^{(\infty)} = \sum_{l=1}^{M}\pi^{(l)}\left(\bar{u}^{(l,\infty)} + \bar{\mathrm{P}}^{\mathrm{load},(l)}\right) - \mathrm{P}^{\mathrm{prod}}$, we obtain using Corollary 4.8 and the definition of $\bar{v}^{(\infty)}$ and $\bar{v}^{(N)}$:

$$\|(u^{(k,i,\infty)} - u^{(k,i,N)}, X^{(k,i,\infty)} - X^{(k,i,N)}, Y^{(k,i,\infty)} - Y^{(k,i,N)})\|_{\mathcal{H}} \le C_T\|\bar{v}^{(\infty)} - \bar{v}^{(N)}\|_{\mathcal{H}} \le \frac{C_T}{\sqrt{N}}.$$

For $k \in [M]$, $i \in [N_k]$, we introduce the notation $u^{\Delta,(k,i)} := u^{(k,i,\infty)} - u^{(k,i,N)}$, for $\sigma \in [0,1]$, $u^{(\sigma),(k,i)} := u^{(k,i,N)} + \sigma(u^{(k,i,\infty)} - u^{(k,i,N)})$, $\bar{u}^{(\sigma)} := \frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k} u^{(\sigma),(k,i)}$.

We have by Taylor formula, by the form of the Gateaux derivative of $\mathcal{J}$ given in (7.2) and since $\dot{\mathcal{J}}(u^{(N)}, u^{(\infty)} - u^{(N)}) = 0$ by optimality of $u^{(N)}$:

$$\begin{aligned}
\mathcal{J}(u^{(\infty)}) - \mathcal{J}(u^{(N)}) &= \int_0^1 \dot{\mathcal{J}}(u^{(N)} + \sigma(u^{(\infty)} - u^{(N)}), u^{(\infty)} - u^{(N)})\mathrm{d}\sigma \\
&= \int_0^1 \left(\dot{\mathcal{J}}(u^{(N)} + \sigma(u^{(\infty)} - u^{(N)}), u^{(\infty)} - u^{(N)}) - \dot{\mathcal{J}}(u^{(N)}, u^{(\infty)} - u^{(N)})\right)\mathrm{d}\sigma \\
&= \int_0^1 \mathbb{E}\left[\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k}\int_0^T \left\{\mu_t^{(k)}\sigma\left(u_t^{(k,i,\infty)} - u_t^{(k,i,N)}\right) + \alpha_t^{(k)}\sigma(Y_t^{(k,i,\infty)} - Y^{(k,i,N)})\right\}(u_t^{(k,i,\infty)} - u_t^{(k,i,N)})\mathrm{d}t\right]\mathrm{d}\sigma \\
&\quad + \int_0^1 \mathbb{E}\left[\int_0^T \left\{\mathcal{L}_x'\left(t, \bar{u}_t^{(\sigma)} + \bar{\mathrm{P}}_t^{\mathrm{load},(N)} - \mathrm{P}_t^{\mathrm{prod}}\right) - \mathcal{L}_x'\left(t, \bar{u}_t^{(0)} + \bar{\mathrm{P}}_t^{\mathrm{load},(N)} - \mathrm{P}_t^{\mathrm{prod}}\right)\right\}\left(\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k} u_t^{\Delta,(k,i)}\right)\mathrm{d}t\right]\mathrm{d}\sigma,
\end{aligned}$$

where we used the affine-linearity of the state and adjoint variables with respect to the control variable.

In what follows, C denotes a constant independent from N, which depends on data of the problem and may change from one line to another.

We then use Cauchy-Schwartz inequality, Taylor formula applied to $\sigma \mapsto \mathcal{L}_x'\left(t, \bar{u}_t^{(\sigma)} + \bar{\mathrm{P}}_t^{\mathrm{load},(N)} - \mathrm{P}_t^{\mathrm{prod}}\right)$ as well as $\bar{u}^{(\sigma)} = \bar{u}^{(0)} + \sigma\left(\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k} u^{\Delta,(k,i)}\right)$ to obtain the following upper bound:

$$\begin{aligned}
\mathcal{J}(u^{(\infty)}) - \mathcal{J}(u^{(N)}) &\le \frac{C}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k}\left(\|Y^{(k,i,N)} - Y^{(k,i,\infty)}\|_{\mathcal{H}} + \|u^{(k,i,N)} - u^{(k,i,\infty)}\|_{\mathcal{H}}\right)\|u^{(k,i,N)} - u^{(k,i,\infty)}\|_{\mathcal{H}} \\
&\quad + \int_0^1 \mathbb{E}\left[\int_0^T\int_0^\sigma \mathcal{L}_{xx}''\left(t, \bar{u}_t^{(r)} + \bar{\mathrm{P}}_t^{\mathrm{load},(N)} - \mathrm{P}_t^{\mathrm{prod}}\right)\left(\frac{1}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k} u_t^{\Delta,(k,i)}\right)^2 \mathrm{d}r\mathrm{d}t\right]\mathrm{d}\sigma.
\end{aligned}$$

Using the boundedness of the second-order derivative of $(t,x) \mapsto \mathcal{L}_{xx}''(t,x)$ uniformly in $t \in [0,T]$ and x, we get:

$$\mathcal{J}(u^{(\infty)}) - \mathcal{J}(u^{(N)}) \le \frac{C}{N}\sum_{k=1}^{M}\sum_{i=1}^{N_k}\left(\|Y^{(k,i,N)} - Y^{(k,i,\infty)}\|_{\mathcal{H}} + \|u^{(k,i,N)} - u^{(k,i,\infty)}\|_{\mathcal{H}}\right)\|u^{(k,i,N)} - u^{(k,i,\infty)}\|_{\mathcal{H}}.$$

Combine this inequality with the previous bound:

$$\|(u^{(k,i,\infty)} - u^{(k,i,N)}, X^{(k,i,\infty)} - X^{(k,i,N)}, Y^{(k,i,\infty)} - Y^{(k,i,N)})\|_{\mathcal{H}} \le \frac{C}{\sqrt{N}},$$

and the fact that $u^{(N)}$ minimizes $\mathcal{J}$ to get:

$$0 \le \mathcal{J}(u^{(\infty)}) - \mathcal{J}(u^{(N)}) \le \frac{C}{N}.$$

8 Conclusion

We have formulated a control problem to model a cooperative setting where Thermostatically Controlled Loads distributed among a large population of agents are used to balance power production and consumption in a context of strong uncertainty created by renewable energy sources. Necessary and sufficient optimality conditions are given, in the form of a high-dimensional FBSDE. The curse of dimensionality one may expect can be dealt with by an appropriate decomposition method, which shows that the high-dimensional FBSDE is equivalent to lower-dimension FBSDEs: a coordination problem and individual problems. In particular, we show the optimal solution of the (centralized) stochastic control problem can be obtained by computing the (unique) Nash equilibrium of an associated Stochastic Stackelberg Differential Game, with a coordinator (leader) solving a control problem and sending coordination signal to the agents (followers), solving their own individual problems. This allows a decentralized implementation. Under a conditional independence-type assumption and in the limit of large population, we show a mean-field type approximation of the problem of the coordinator, which does not require the aggregator to observe the behaviors of the agents, in the framework of the associated Stochastic differential game. This is desirable for both preserving privacy of consumers and reducing the need for real-time telecommunication between agents and the coordinator, as in a federated learning paradigm. Numerical results show the performance of the approach and the quality of the mean-field approximation. The experiments demonstrate the need to carefully tune the cost parameters of the problem in order to maximize the contribution of the TCLs to power balancing while ensuring that individual constraints of the devices are not violated. A decentralized and online implementation of the control mechanism with minimal one-way communication from the aggregator to the agents is also proposed, allowing for the coordinator and the agents to solve in real-time their respective problems.

References

[AA19] A. Agarkar and H. Agrawal. "A review and vision on authentication and privacy preservation schemes in smart grid network". In: *Security and Privacy* 2.2 (2019), e62.

[ATM20] C. Alasseur, I. B. Taher, and A. Matoussi. "An extended mean field game for storage in smart grids". In: *Journal of Optimization Theory and Applications* 184.2 (2020), pp. 644–670.

[ADS99] M. Alexiadis, P. Dokopoulos, and H. Sahsamanoglou. "Wind speed and power forecasting based on spatial correlation models". In: *IEEE Transactions on Energy Conversion* 14.3 (1999), pp. 836–842.

[Bad+18] J. Badosa, E. Gobet, M. Grangereau, and D. Kim. "Day-ahead probabilistic forecast of solar irradiance: a Stochastic Differential Equation approach". In: *Renewable Energy: Forecasting and Risk Management*. Ed. by P. Drobinski, M. Mougeot, D. Picard, R. Plougonven, and P. Tankov. Springer Proceedings in Mathematics & Statistics, 2018. Chap. 4, pp. 73–93.

[BFY+13] A. Bensoussan, J. Frehse, P. Yam, et al. *Mean field games and mean field type control theory*. Vol. 101. Springer, 2013.

[BSS05] B. Bibby, I. Skovgaard, and M. Sorensen. "Diffusion-type models with given marginal distribution and autocorrelation function". In: *Bernoulli* 11.2 (2005), pp. 191–220.

[BG19] R. Binns and V. Gallo. "Data minimisation and privacy-preserving techniques in AI systems". In: *UK Information Commissioners Office* (2019). URL: https://ico.org.uk/about-the-ico/news-and-events/ai-blog-data-minimisation-and-privacy-preserving-techniques-in-ai-systems/.

[Bis76] J.-M. Bismut. "Linear quadratic optimal stochastic control with random coefficients". In: *SIAM Journal on Control and Optimization* 14.3 (1976), pp. 419–444.

[Bre10] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer New York, 2010. ISBN: 9780387709130.

[BM16] A. Bušić and S. Meyn. "Distributed randomized control for demand dispatch". In: *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE. 2016, pp. 6964–6971.

[Cad02] A. Cadenillas. “A stochastic maximum principle for systems with jumps, with applications to finance”. In: *Systems & control letters* 47.5 (2002), pp. 433–444.

[Cam+18] N. J. Cammardella, R. W. Moye, Y. Chen, and S. P. Meyn. “An energy storage cost comparison: Li-ion batteries vs distributed load control”. In: *2018 Clemson University Power Systems Conference (PSC)*. IEEE. 2018, pp. 1–6.

[CD18] R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications I-II*. Springer, 2018.

[CDL13] R. Carmona, F. Delarue, and A. Lachapelle. “Control of McKean-Vlasov dynamics versus mean field games”. In: *Math. Financ. Econ.* 7.2 (2013), pp. 131–166.

[Car+19] P. Carpentier, J.-P. Chancelier, M. de Lara, and F. Pacaud. “Upper and lower bounds for large scale multistage stochastic optimization problems: Decomposition methods”. In: *arXiv preprint arXiv:1912.10901* (2019).

[Car+20] P. Carpentier, J.-P. Chancelier, M. De Lara, and F. Pacaud. “Mixed Spatial and Temporal Decompositions for Large-Scale Multistage Stochastic Optimization Problems”. In: *Journal of Optimization Theory and Applications* 186.3 (2020), pp. 985–1005.

[DP+19] A. De Paola, V. Trovato, D. Angeli, and G. Strbac. “A mean field game approach for distributed control of thermostatic loads acting in simultaneous energy-frequency response markets”. In: *IEEE Transactions on Smart Grid* 10.6 (2019), pp. 5987–5999.

[EPQ97] N. El Karoui, S. Peng, and M. Quenez. “Backward stochastic differential equations in finance”. In: *Math. Finance* 7.1 (1997), pp. 1–71.

[FMHL19] A Fonseca-Morales and O Hernández-Lerma. “Stochastic differential games: the potential approach”. In: *Stochastics* (2019), pp. 1–14.

[Hao+13] H. Hao, B. M. Sanandaji, K. Poolla, and T. L. Vincent. “A generalized battery model of a collection of thermostatically controlled loads for providing ancillary service”. In: *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE. 2013, pp. 551–558.

[Hao+14] H. Hao, B. M. Sanandaji, K. Poolla, and T. L. Vincent. “Aggregate flexibility of thermostatically controlled loads”. In: *IEEE Transactions on Power Systems* 30.1 (2014), pp. 189–198.

[HMC+06] M. Huang, R. P. Malhamé, P. E. Caines, et al. “Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle”. In: *Communications in Information & Systems* 6.3 (2006), pp. 221–252.

[Jac+19] P. Jacquot, O. Beaude, P. Benchimol, S. Gaubert, and N. Oudjane. “A privacy-preserving disaggregation algorithm for non-intrusive management of flexible energy”. In: *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE. 2019, pp. 890–896.

[KS98] I. Karatzas and S. E. Shreve. “Brownian motion”. In: *Brownian Motion and Stochastic Calculus*. Springer, 1998, pp. 47–127.

[KM13] A. C. Kizilkale and R. P. Malhamé. “Mean field based control of power system dispersed energy storage devices for peak load relief”. In: *52nd IEEE Conference on Decision and Control*. IEEE. 2013, pp. 4971–4976.

[KM16] A. C. Kizilkale and R. P. Malhamé. “Collective target tracking mean field control for Markovian jump-driven models of electric water heating loads”. In: *Control of Complex Systems*. Elsevier, 2016, pp. 559–584.

[Ma+15] J. Ma, Z. Wu, D. Zhang, J. Zhang, et al. “On well-posedness of forward–backward SDEs—A unified approach”. In: *The Annals of Applied Probability* 25.4 (2015), pp. 2168–2214.

[Mat+12] J Mathieu, M. Dyson, D. Callaway, and A Rosenfeld. “Using residential electric loads for fast demand response: The potential resource and revenues, the costs, and policy recommendations”. In: *ACEEE Summer Study on Energy Efficiency in Buildings*. Citeseer. 2012, pp. 189–203.

[MMS19] A. Matoussi, A. Manai, and R. Salhi. “Mean-Field Backward-Forward SDE with Jumps and Storage problem in Smart Grids”. In: *arXiv preprint arXiv:1906.08525* (2019).

[PMV02] A. Pardo, V. Meneu, and E. Valor. “Temperature and seasonality influences on Spanish electricity load”. In: *Energy Economics* 24.1 (2002), pp. 55–70.

[Pro03] P. E. Protter. *Stochastic Integration and Differential Equations*. second. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003. ISBN: 978-3-662-10061-5. DOI: 10.1007/978-3-662-10061-5_6. URL: https://doi.org/10.1007/978-3-662-10061-5_6.

[ŞC14] N. Şen and P. E. Caines. “Mean field games with partially observed major player and stochastic mean field”. In: *53rd IEEE Conference on Decision and Control*. IEEE. 2014, pp. 2709–2715.

[TTS15] S. H. Tindemans, V. Trovato, and G. Strbac. “Decentralized control of thermostatic loads for flexible demand response”. In: *IEEE Transactions on Control Systems Technology* 23.5 (2015), pp. 1685–1700.

[TTS16] V. Trovato, S. H. Tindemans, and G. Strbac. “Leaky storage model for optimal multi-service allocation of thermostatic loads”. In: *IET Generation, Transmission & Distribution* 10.3 (2016), pp. 585–593.

[Tro+16] V. Trovato, I. M. Sanz, B. Chaudhuri, and G. Strbac. “Advanced control of thermostatic loads for rapid frequency response in Great Britain”. In: *IEEE Transactions on Power Systems* 32.3 (2016), pp. 2106–2117.

[YFY20] Q. Yang, L. Fan, and H. Yu, eds. *Federated Learning. Privacy and incentive*. Vol. 12500. Lecture Notes in Artificial Intelligence. Springer Nature Switzerland, 2020.

[Yao86] A. C.-C. Yao. “How to generate and exchange secrets”. In: *27th Annual Symposium on Foundations of Computer Science (sfcs 1986)*. IEEE. 1986, pp. 162–167.

[Yon06] J. Yong. “Linear forward-backward stochastic differential equations with random coefficients”. In: *Probability theory and related fields* 135.1 (2006), pp. 53–83.

[Yon99] J. Yong. “Linear Forward—Backward Stochastic Differential Equations”. In: *Applied Mathematics and Optimization* 39.1 (1999), pp. 93–119.

[ZDK16] B. Zhang, P. Dehghanian, and M. Kezunovic. “Spatial-temporal solar power forecast through use of Gaussian conditional random fields”. In: *2016 IEEE Power and Energy Society General Meeting (PESGM)*. IEEE. 2016, pp. 1–5.